(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

10/018170

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2000 (21.12.2000)

PCT

(10) International Publication Number
WO 00/077040 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, C12Q 1/68, A61K 38/17

(21) International Application Number: PCT/US00/16636

(22) International Filing Date: 16 June 2000 (16.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/139,566	16 June 1999 (16.06.1999)	US
60/149,640	17 August 1999 (17.08.1999)	US
60/164,417	9 November 1999 (09.11.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/149,640 (CIP)
Filed on	17 August 1999 (17.08.1999)
US	60/164,417 (CIP)
Filed on	9 November 1999 (09.11.1999)
US	60/139,566 (CIP)
Filed on	16 June 1999 (16.06.1999)

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). **LAL, Preeti** [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **AZIMZAI, Yalda** [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). **YANG, Junming** [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). **REDDY, Roopa** [IN/US]; 1233 W. McKinley Avenue #3, Sunnyvale, CA 94086 (US). **LU, Dyung, Aina, M.** [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report: 26 April 2001

(48) Date of publication of this corrected version: 1 August 2002

(15) Information about Correction:
see PCT Gazette No. 31/2002 of 1 August 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/077040 A3

(54) Title: HUMAN INTRACELLULAR SIGNALING MOLECULES

(57) Abstract: The invention provides human intracellular signaling molecules (INTRA) and polynucleotides which identify and encode INTRA. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of INTRA.

INTRACELLULAR SIGNALING MOLECULES

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of intracellular signaling molecules and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, neurological, gastrointestinal, reproductive, and developmental disorders.

BACKGROUND OF THE INVENTION

Cell-cell communication is essential for the growth, development, and survival of multicellular organisms. Cells communicate by sending and receiving molecular signals. An example of a molecular signal is a growth factor, which binds and activates a specific transmembrane receptor on the surface of a target cell. The activated receptor transduces the signal intracellularly, thus initiating a cascade of biochemical reactions that ultimately affect gene transcription and cell cycle progression in the target cell.

Intracellular signaling is the process by which cells respond to extracellular signals (hormones, neurotransmitters, growth and differentiation factors, etc.) through a cascade of biochemical reactions that begins with the binding of a signaling molecule to a cell membrane receptor and ends with the activation of an intracellular target molecule. Intermediate steps in the process involve the activation of various cytoplasmic proteins by phosphorylation via protein kinases, and their deactivation by protein phosphatases, and the eventual translocation of some of these activated proteins to the cell nucleus where the transcription of specific genes is triggered. The intracellular signaling process regulates all types of cell functions including cell proliferation, cell differentiation, and gene transcription, and involves a diversity of molecules including protein kinases and phosphatases, and second messenger molecules such as cyclic nucleotides, calcium-calmodulin, inositol, and various mitogens that regulate protein phosphorylation.

Intracellular signaling is carried out by a variety of molecules that promote the transduction and amplification of the signal. For example, binding of a ligand to a transmembrane receptor activates membrane-associated intracellular proteins, such as G-proteins. G-proteins mediate both the level of intracellular second messengers, such as cyclic AMP, and the activity of signaling enzymes, such as phospholipase C. These messengers and enzymes then activate signal transduction pathways, many of which are mediated by protein kinase cascades. Phosphorylation of proteins in response to extracellular signals, cell cycle checkpoints, and environmental or nutritional stresses is often accomplished by transfer of a high energy phosphate from ATP. Second messengers whose effects are mediated by protein kinases include cyclic AMP, cyclic GMP, inositol triphosphate, cyclic ADP

ribose, and calcium/calmodulin. Alternatively, binding of ligand to a transmembrane receptor, such as a receptor tyrosine kinase, triggers the activation of a molecular "switch," such as a monomeric GTPase. In this case, binding of ligand to the receptor activates a catalytic domain in the intracellular portion of the receptor. This activated domain then switches on the activity of monomeric GTPases such as Ras, usually via adaptor proteins.

Cells also respond to changing conditions by switching off signals. Many signal transduction proteins are short-lived and rapidly targeted for degradation by covalent ligation to ubiquitin, a highly conserved small protein. Cells also maintain mechanisms to monitor changes in the concentration of denatured or unfolded proteins in membrane-bound extracytoplasmic compartments, including a transmembrane receptor that monitors the concentration of available chaperone molecules in the endoplasmic reticulum and transmits a signal to the cytosol to activate the transcription of nuclear genes encoding chaperones in the endoplasmic reticulum.

Certain proteins in intracellular signaling pathways serve to link or cluster other proteins involved in the signaling cascade. These proteins are referred to as scaffold, anchoring, or adaptor proteins. (For review, see Pawson, T., and Scott, J.D. (1997) Science 278:2075-2080.) As many intracellular signaling proteins such as protein kinases and phosphatases have relatively broad substrate specificities, the adaptors help to organize the component signaling proteins into specific biocehmical pathways.

Gangliosides, generally associated with plasma membranes, also participate in signal transduction. Aberrant ganglioside function has been implicated in inflammatory and degenerative diseases within and outside of the nervous system, including Tay-Sachs disease, multiple sclerosis, lupus erythematosus, and insulin-dependent diabetes mellitus (Misasi, R. et al. (1997) Diabetes Metab. Rev. 13:163-179).

Many of the above signaling molecules are characterized by the presence of particular domains that promote protein-protein interactions. A sampling of these domains is discussed below, along with other important intracellular messengers.

Intracellular Signaling Second Messenger Molecules

Phospholipid and Inositol-phosphate Signaling

Inositol phospholipids (phosphoinositides) are involved in an intracellular signaling pathway that begins with binding of a signaling molecule to a G-protein linked receptor in the plasma membrane. This leads to the phosphorylation of phosphatidylinositol (PI) residues on the inner side of the plasma membrane to the biphosphate state (PIP_2) by inositol kinases. Simultaneously, the G-protein linked receptor binding stimulates a trimeric G-protein which in turn activates a phosphoinositide-specific phospholipase C-β. Phospholipase C-β then cleaves PIP_2 into two

products, inositol triphosphate (IP_3) and diacylglycerol. These two products act as mediators for separate signaling events. IP_3 diffuses through the plasma membrane to induce calcium release from the endoplasmic reticulum (ER), while diaacylglycerol remains in the membrane and helps activate protein kinase C, an STK that phosphorylates selected proteins in the target cell. The calcium response initiated by IP_3 is terminated by the dephosphorylation of IP_3 by specific inositol phosphatases. Cellular responses that are mediated by this pathway are glycogen breakdown in the liver in response to vasopressin, smooth muscle contraction in response to acetylcholine, and thrombin-induced platelet aggregation.

Cyclic Nucleotide Signaling

Cyclic nucleotides (cAMP and cGMP) function as intracellular second messengers to transduce a variety of extracellular signals including hormones, light, and neurotransmitters. In particular , cyclic-AMP dependent protein kinases (PKA) are thought to account for all of the effects of cAMP in most mammalian cells, including various hormone-induced cellular responses. Visual excitation and the phototransmission of light signals in the eye is controlled by cyclic-GMP regulated, Ca^{2+}-specific channels. Because of the importance of cellular levels of cyclic nucleotides in mediating these various responses, regulating the synthesis and breakdown of cyclic nucleotides is an important matter. Thus adenylyl cyclase, which synthesizes cAMP from AMP, is activated to increase cAMP levels in muscle by binding of adrenaline to β-andrenergic receptors, while activation of guanylate cyclase and increased cGMP levels in photoreceptors leads to reopening of the Ca^{2+}-specific channels and recovery of the dark state in the eye. In contrast, hydrolysis of cyclic nucleotides by cAMP and cGMP-specific phosphodiesterases (PDEs) produces the opposite of these and other effects mediated by increased cyclic nucleotide levels. PDEs appear to be particularly important in the regulation of cyclic nucleotides, considering the diversity found in this family of proteins. At least seven families of mammalian PDEs (PDE1-7) have been identified based on substrate specificity and affinity, sensitivity to cofactors, and sensitivity to inhibitory drugs (Beavo, J.A. (1995) Physiological Reviews 75:725-48). PDE inhibitors have been found to be particularly useful in treating various clinical disorders. Rolipram, a specific inhibitor of PDE4, has been used in the treatment of depression, and similar inhibitors are undergoing evaluation as anti-inflammatory agents. Theophylline is a nonspecific PDE inhibitor used in the treatment of bronchial asthma and other respiratory diseases (Banner, K.H. and Page, C.P. (1995) Eur. Respir. J. 8:996-1000).

Calcium Signaling Molecules

Ca^{+2} is another second messenger molecule that is even more widely used as an intracellular mediator than cAMP. Two pathways exist by which Ca^{+2} can enter the cytosol in response to extracellular signals: One pathway acts primarily in nerve signal transduction where Ca^{+2} enters a nerve terminal through a voltage-gated Ca^{+2} channel. The second is a more ubiquitous pathway in

which Ca^{+2} is released from the ER into the cytosol in response to binding of an extracellular signaling molecule to a receptor. Ca^{2+} directly activates regulatory enzymes, such as protein kinase C, which trigger signal transduction pathways. Ca^{2+} also binds to specific Ca^{2+}-binding proteins (CBPs) such as calmodulin (CaM) which then activate multiple target proteins in the cell including enzymes, membrane transport pumps, and ion channels. CaM interactions are involved in a multitude of cellular processes including, but not limited to, gene regulation, DNA synthesis, cell cycle progression, mitosis, cytokinesis, cytoskeletal organization, muscle contraction, signal transduction, ion homeostasis, exocytosis, and metabolic regulation (Celio, M.R. et al. (1996) Guidebook to Calcium-binding Proteins, Oxford University Press, Oxford, UK, pp. 15-20). Some Ca^{2+} binding proteins are characterized by the presence of one or more EF-hand Ca^{2+} binding motifs, which are comprised of 12 amino acids flanked by α-helices (Celio, supra). The regulation of CBPs has implications for the control of a variety of disorders. Calcineurin, a CaM-regulated protein phosphatase, is a target for inhibition by the immunosuppressive agents cyclosporin and FK506. This indicates the importance of calcineurin and CaM in the immune response and immune disorders (Schwaninger M. et al. (1993) J. Biol Chem. 268:23111-23115). The level of CaM is increased several-fold in tumors and tumor-derived cell lines for various types of cancer (Rasmussen, C.D. and Means, A.R. (1989) Trends in Neuroscience 12:433-438).

The annexins are a family of calcium-binding proteins that associate with the cell membrane (Towle, C.A. and Treadwell, B.V. (1992) J. Biol. Chem. 267:5416-23). Annexins reversibly bind to negatively charged phospholipids (phosphatidylcholine and phosphatidylserine) in a calcium dependent manner. Annexins participate in various processes pertaining to signal transduction at the plasma membrane, including membrane-cytoskeleton interactions, phospholipase inhibition, anticoagulation, and membrane fusion. Annexins contain four to eight repeated segments of about 60 residues. Each repeat folds into five alpha helices wound into a right-handed superhelix.

Signaling Complex Protein Domains

PDZ domains were named for three proteins in which this domain was initially discovered. These proteins include PSD-95 (postsynaptic density 95), Dlg (Drosophila lethal(1)discs large-1), and ZO-1 (zonula occludens-1). These proteins play important roles in neuronal synaptic transmission, tumor suppression, and cell junction formation, respectively. Since the discovery of these proteins, over sixty additional PDZ-containing proteins have been identified in diverse prokaryotic and eukaryotic organisms. This domain has been implicated in receptor and ion channel clustering and in the targeting of multiprotein signaling complexes to specialized functional regions of the cytosolic face of the plasma membrane. (For review of PDZ domain-containing proteins, see Ponting, C. P. et al. (1997) Bioessays 19:469-479.) A large proportion of PDZ domains are found in the eukaryotic MAGUK (membrane-associated guanylate kinase) protein family, members of which bind to the

intracellular domains of receptors and channels. However, PDZ domains are also found in diverse membrane-localized proteins such as protein tyrosine phosphatases, serine/threonine kinases, G-protein cofactors, and synapse-associated proteins such as syntrophins and neuronal nitric oxide synthase (nNOS). Generally, about one to three PDZ domains are found in a given protein, although up to nine PDZ domains have been identified in a single protein. The glutamate receptor interacting protein (GRIP) contains seven PDZ domains. GRIP is an adaptor that links certain glutamate receptors to other proteins and may be responsible for the clustering of these receptors at excitatory synapses in the brain (Dong, H. et al. (1997) Nature 386:279-284).

The SH3 domain is defined by homology to a region of the proto-oncogene c-Src, a cytoplasmic protein tyrosine kinase. SH3 is a small domain of 50 to 60 amino acids that interacts with proline-rich ligands. SH3 domains are found in a variety of eukaryotic proteins involved in signal transduction, cell polarization, and membrane-cytoskeleton interactions. In some cases, SH3 domain-containing proteins interact directly with receptor tyrosine kinases. For example, the SLAP-130 protein is a substrate of the T-cell receptor (TCR) stimulated protein kinase. SLAP-130 interacts via its SH3 domain with the protein SLP-76 to affect the TCR-induced expression of interleukin-2 (Musci, M.A. et al. (1997) J. Biol. Chem. 272:11674-11677). Another recently identified SH3 domain protein is macrophage actin-associated tyrosine-phosphorylated protein (MAYP) which is phosphorylated during the response of macrophages to colony stimulating factor-1 (CSF-1) and is likely to play a role in regulating the CSF-1-induced reorganization of the actin cytoskeleton (Yeung, Y.-G. et al. (1998) J. Biol. Chem. 273:30638-30642). The structure of SH3 is characterized by two antiparallel beta sheets packed against each other at right angles. This packing forms a hydrophobic pocket lined with residues that are highly conserved between different SH3 domains. This pocket makes critical hydrophobic contacts with proline residues in the ligand (Feng, S. et al. (1994) Science 266: 1241-47). Endophilin is an SH3 domain-containing protein implicated in synaptic vesicle endocytosis. (Micheva, K.D. (1997) 272:27239-27245).

A novel domain, called the WW domain, resembles the SH3 domain in its ability to bind proline-rich ligands. This domain was originally discovered in dystrophin, a cytoskeletal protein with direct involvement in Duchenne muscular dystrophy (Bork, P. and Sudol, M. (1994) Trends Biochem. Sci. 19:531-533). WW domains have since been discovered in a variety of intracellular signaling molecules involved in development, cell differentiation, and cell proliferation. The structure of the WW domain is composed of beta strands grouped around four conserved aromatic residues, generally tryptophan.

Like SH3, the SH2 domain is defined by homology to a region of c-Src. SH2 domains interact directly with phospho-tyrosine residues, thus providing an immediate mechanism for the regulation and transduction of receptor tyrosine kinase-mediated signaling pathways. For example, as

many as ten distinct SH2 domains are capable of binding to phosphorylated tyrosine residues in the activated PDGF receptor, thereby providing a highly coordinated and finely tuned response to ligand-mediated receptor activation. (Reviewed in Schaffhausen, B. (1995) Biochem. Biophys. Acta. 1242:61-75.)

Homer is a neuronal immediate early gene that is enriched at excitatory synapses (Xiao, B. et al. (1998) Neuron 21:707-716). Homer proteins form multivalent complexes that bind proline-rich motifs in group 1 metabotropic glutamate receptors and inositol triphosphate receptors, thereby coupling these receptors in a signaling complex (Tu, J.C. (1999) Neuron 23:583-592).

The pleckstrin homology (PH) domain was originally identified in pleckstrin, the predominant substrate for protein kinase C in platelets. Since its discovery, this domain has been identified in over 90 proteins involved in intracellular signaling or cytoskeletal organization. Proteins containing the pleckstrin homology domain include a variety of kinases, phospholipase-C isoforms, guanine nucleotide release factors, and GTPase activating proteins. For example, members of the FGD1 family contain both Rho-guanine nucleotide exchange factor (GEF) and PH domains, as well as a FYVE zinc finger domain. FGD1 is the gene responsible for faciogenital dysplasia, an inherited skeletal dysplasia (Pasteris, N.G. and Gorski, J.L. (1999) Genomics 60:57-66). Many PH domain proteins function in association with the plasma membrane, and this association appears to be mediated by the PH domain itself. PH domains share a common structure composed of two antiparallel beta sheets flanked by an amphipathic alpha helix. Variable loops connecting the component beta strands generally occur within a positively charged environment and may function as ligand binding sites (Lemmon, M. A. et al. (1996) Cell 85:621-624.). n-Chimaerin is a GAP involved in the formation of lamellipodia and filopodia in neuroblastoma cells. (Kozma, R. et al. (1996) Mol. Cell Biol. 16:5069-5080.)

Ankyrin (ANK) repeats mediate protein-protein interactions associated with diverse intracellular signaling functions. For example, ANK repeats are found in proteins involved in cell proliferation such as kinases, kinase inhibitors, tumor suppressors, and cell cycle control proteins. (See, for example, Kalus, W. et al. (1997) FEBS Lett. 401:127-132; Ferrante, A. W. et al. (1995) Proc. Natl. Acad. Sci. USA 92:1911-1915.) These proteins generally contain multiple ANK repeats, each composed of about 33 amino acids. Myotrophin is an ANK repeat protein that plays a key role in the development of cardiac hypertrophy, a contributing factor to many heart diseases. Structural studies show that the myotrophin ANK repeats, like other ANK repeats, each form a helix-turn-helix core preceded by a protruding "tip." These tips are of variable sequence and may play a role in protein-protein interactions. The helix-turn-helix region of the ANK repeats stack on top of one another and are stabilized by hydrophobic interactions (Yang, Y. et al. (1998) Structure 6:619-626).

The tetratrico peptide repeat (TPR) is a 34 amino acid repeated motif found in organisms

from bacteria to humans. TPRs are predicted to form ampipathic helices, and appear to mediate protein-protein interactions. TPR domains are found in CDC16, CDC23, and CDC27, members the the anaphase promoting complex which targets proteins for degradation at the onset of anaphase. Other processes involving TPR proteins include cell cycle control, transcription repression, stress response, and protein kinase inhibition. (Lamb, J.R. et al. (1995) Trends Biochem. Sci. 20:257-259.)

The armadillo/beta-catenin repeat is a 42 amino acid motif which forms a superhelix of alpha helices when tandemly repeated. The structure of the armadillo repeat region from beta-catenin revealed a shallow groove of positive charge on one face of the superhelix, which is a potential binding surface. The armadillo repeats of beta-catenin, plakoglobin, and p120^{cas} bind the cytoplasmic domains of cadherins. Beta-catenin/cadherin complexes are targets of regulatory signals that govern cell adhesion and mobility. (Huber, A.H. et al. (1997) Cell 90:871-882.)

The discovery of new intracellular signaling proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, reproductive, and developmental disorders.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, intracellular signaling molecules, referred to collectively as "INTRA" and individually as "INTRA-1," "INTRA-2," "INTRA-3," "INTRA-4," "INTRA-5," "INTRA-6," "INTRA-7," "INTRA-8," "INTRA-9," "INTRA-10," "INTRA-11," "INTRA-12," "INTRA-13," "INTRA-14," "INTRA-15," "INTRA-16," "INTRA-17," "INTRA-18," "INTRA-19," "INTRA-20," "INTRA-21," "INTRA-22," "INTRA-23," "INTRA-24," "INTRA-25," "INTRA-26," "INTRA-27," "INTRA-28," "INTRA-29," "INTRA-30," "INTRA-31," "INTRA-32," "INTRA-33," "INTRA-34," "INTRA-35," "INTRA-36," "INTRA-37," "INTRA-38," "INTRA-39," "INTRA-40," "INTRA-41," "INTRA-42," "INTRA-43," "INTRA-44," "INTRA-45," "INTRA-46," "INTRA-47," "INTRA-48," "INTRA-49," "INTRA-50," "INTRA-51," and "INTRA-52." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-52.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising

an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-52. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:53-104.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence

selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and a pharmaceutically acceptable excipient. In one embodiment, the pharmaceutical composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional INTRA, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional INTRA, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional INTRA, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an

amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-52. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:53-104, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding INTRA.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of INTRA.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases,

disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding INTRA were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"INTRA" refers to the amino acid sequences of substantially purified INTRA obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of INTRA. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of INTRA either by directly interacting with INTRA or by acting on components of the biological pathway in which INTRA participates.

An "allelic variant" is an alternative form of the gene encoding INTRA. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding INTRA include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as INTRA or a polypeptide with at least one functional characteristic of INTRA. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding INTRA, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding INTRA. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent INTRA. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of INTRA is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of INTRA. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of INTRA either by directly interacting with INTRA or by acting on components of the biological pathway in which

INTRA participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind INTRA polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic INTRA, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding INTRA or fragments of INTRA may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide

backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of INTRA or the polynucleotide encoding INTRA which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:53-104 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:53-104, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:53-104 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:53-104 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:53-104 and the region of SEQ ID NO:53-104 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-52 is encoded by a fragment of SEQ ID NO:53-104. A fragment of SEQ ID NO:1-52 comprises a region of unique amino acid sequence that specifically

identifies SEQ ID NO:1-52. For example, a fragment of SEQ ID NO:1-52 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-52. The precise length of a fragment of SEQ ID NO:1-52 and the region of SEQ ID NO:1-52 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to

compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise

comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature

under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of INTRA which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of INTRA which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of INTRA. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of INTRA.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an INTRA may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of INTRA.

"Probe" refers to nucleic acid sequences encoding INTRA, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also

be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2[nd] ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence

that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding INTRA, or fragments thereof, or INTRA itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are

removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of

the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human intracellular signaling molecules (INTRA), the polynucleotides encoding INTRA, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, neurological, gastrointestinal, reproductive, and developmental disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding INTRA. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each INTRA were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each INTRA and are useful as fragments in

hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis along with relevant citations, all of which are expressly incorporated by reference herein in their entirety; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding INTRA. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:53-104 and to distinguish between SEQ ID NO:53-104 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express INTRA as a fraction of total tissues expressing INTRA. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing INTRA as a fraction of total tissues expressing INTRA. Column 5 lists the vectors used to subclone each cDNA library. Of particular interest is the expression of SEQ ID NO:88 and SEQ ID NO:94 in reproductive tissues, of SEQ ID NO:99, SEQ ID NO:100, and SEQ ID NO:103 in hematopoietic/immune tissues, and of SEQ ID NO:96 in cardiovascular tissues.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding INTRA were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:58 maps to chromosome 7 within the interval from 84.40 to 90.30 centiMorgans. This interval also contains an EST with high similarity to thyroid disease hypothetical autoantigen. SEQ ID NO:67 maps to chromosome 16 within the interval from 119.20 centiMorgans to q-terminus. This interval also contains the paraplegin gene, mutations in which cause spastic paraplegia and OXPHOS impairment. SEQ ID NO:70 maps to chromosome 11 within the interval from 59.50 to 62.50 centiMorgans. SEQ ID NO:71 maps to chromosome 7 within the interval from 138.0 to 145.8 centiMorgans. SEQ ID NO:73 maps to chromosome 12 within the interval from 76.5 to 84.2 centiMorgans. SEQ ID NO:77 maps to chromosome 7 within the interval from 4.8 to 10.6 centiMorgans and to chromosome 4 within the interval from 56.7 to 60.5 centiMorgans. The interval

on chromosome 7 from from 4.8 to 10.6 centiMorgans also contains a gene associated with cell proliferation. The interval on chromosome 4 from 56.7 to 60.5 centiMorgans also contains a gene associated with cell proliferation. SEQ ID NO:79 maps to chromosome 15 within the interval from 32.2 to 47.1 centiMorgans. This interval also contains a gene associated with cell proliferation. SEQ ID NO:80 maps to chromosome 20 within the interval from 50.2 to 53.6 centiMorgans. This interval also contains a gene associated with cell differentiation. SEQ ID NO:84 maps to chromosome 3 within the interval from 142.2 to 148.7 centiMorgans. SEQ ID NO:87 maps to chromosome 5 within the interval from 141.4 to 147.1 centiMorgans. SEQ ID NO:91 maps to chromosome 12 within the interval from 62.7 to 67.3 centiMorgans. SEQ ID NO:95 maps to chromosome 15 within the interval from 45.5 to 58.8 centiMorgans. SEQ ID NO:97 maps to the X chromosome within the interval from 112.8 to 139.4 centiMorgans.

The invention also encompasses INTRA variants. A preferred INTRA variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the INTRA amino acid sequence, and which contains at least one functional or structural characteristic of INTRA.

The invention also encompasses polynucleotides which encode INTRA. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:53-104, which encodes INTRA. The polynucleotide sequences of SEQ ID NO:53-104, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding INTRA. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding INTRA. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:53-104 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:53-104. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of INTRA.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding INTRA, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These

combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring INTRA, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode INTRA and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring INTRA under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding INTRA or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding INTRA and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode INTRA and INTRA derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding INTRA or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:53-104 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting

sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding INTRA may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate

software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode INTRA may be cloned in recombinant DNA molecules that direct expression of INTRA, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express INTRA.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter INTRA-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of INTRA, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding INTRA may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, INTRA itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of INTRA, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active INTRA, the nucleotide sequences encoding INTRA or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding INTRA. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding INTRA. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding INTRA and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding INTRA and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding INTRA. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al., (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding INTRA. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding INTRA can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding INTRA into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of INTRA are needed, e.g. for the production of antibodies, vectors which direct high level expression of INTRA may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of INTRA. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH

promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of INTRA. Transcription of sequences encoding INTRA may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding INTRA may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses INTRA in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of INTRA in cell lines is preferred. For example, sequences encoding INTRA can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding INTRA is inserted within a marker gene sequence, transformed cells containing sequences encoding INTRA can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding INTRA under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding INTRA and that express INTRA may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of INTRA using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on INTRA is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and

Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding INTRA include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding INTRA, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding INTRA may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode INTRA may be designed to contain signal sequences which direct secretion of INTRA through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding INTRA may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric INTRA protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of INTRA activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available

affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the INTRA encoding sequence and the heterologous protein sequence, so that INTRA may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled INTRA may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

INTRA of the present invention or fragments thereof may be used to screen for compounds that specifically bind to INTRA. At least one and up to a plurality of test compounds may be screened for specific binding to INTRA. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of INTRA, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which INTRA binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express INTRA, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing INTRA or cell membrane fractions which contain INTRA are then contacted with a test compound and binding, stimulation, or inhibition of activity of either INTRA or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with INTRA, either in

solution or affixed to a solid support, and detecting the binding of INTRA to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

INTRA of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of INTRA. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for INTRA activity, wherein INTRA is combined with at least one test compound, and the activity of INTRA in the presence of a test compound is compared with the activity of INTRA in the absence of the test compound. A change in the activity of INTRA in the presence of the test compound is indicative of a compound that modulates the activity of INTRA. Alternatively, a test compound is combined with an in vitro or cell-free system comprising INTRA under conditions suitable for INTRA activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of INTRA may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding INTRA or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding INTRA may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate

into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding INTRA can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding INTRA is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress INTRA, e.g., by secreting INTRA in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of INTRA and intracellular signaling molecules. In addition, the expression of INTRA is closely associated with cancers of the hematopoetic/immune, nervous, gastrointestinal, and reproductive, systems therefore, INTRA appears to play a role in cell proliferative, autoimmune/inflammatory, neurological, gastrointestinal, reproductive, and developmental disorders. In the treatment of disorders associated with increased INTRA expression or activity, it is desirable to decrease the expression or activity of INTRA. In the treatment of disorders associated with decreased INTRA expression or activity, it is desirable to increase the expression or activity of INTRA.

Therefore, in one embodiment, INTRA or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of INTRA. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, hematopoietic cancer including lymphoma, leukemia, and myeloma; and other cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, adenoma, carcinoma and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's

disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and a hepatic tumor including a nodular hyperplasia,.a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system

disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and a gastrointestinal disorder such as esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease, cholecystitis, infections of the intestinal tract, pancreatitis, pancreatic carcinoma, cirrhosis, hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease.

In another embodiment, a vector capable of expressing INTRA or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of INTRA including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified INTRA in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of INTRA including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of INTRA may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of INTRA including, but not limited to, those listed above.

In a further embodiment, an antagonist of INTRA may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of INTRA. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, neurological, gastrointestinal, reproductive, and developmental disorders described above. In one aspect, an antibody which specifically binds INTRA may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express INTRA.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding INTRA may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of INTRA including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of INTRA may be produced using methods which are generally known in the art. In particular, purified INTRA may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind INTRA. Antibodies to INTRA may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with INTRA or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to INTRA have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of INTRA amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to INTRA may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce INTRA-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for INTRA may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between INTRA and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering INTRA epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for INTRA. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of INTRA-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple INTRA epitopes, represents the average affinity, or avidity, of the antibodies for INTRA. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular INTRA epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the INTRA-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of INTRA, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml,

preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of INTRA-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding INTRA, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding INTRA. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding INTRA. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding INTRA may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated

cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in INTRA expression or regulation causes disease, the expression of INTRA from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in INTRA are treated by constructing mammalian expression vectors encoding INTRA and introducing these vectors by mechanical means into INTRA-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of INTRA include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). INTRA may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding INTRA from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al.

(1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to INTRA expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding INTRA under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. U.S.A. 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding INTRA to cells which have one or more genetic abnormalities with respect to the expression of INTRA. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding INTRA to target cells which have one or more genetic abnormalities with

respect to the expression of INTRA. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing INTRA to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding INTRA to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for INTRA into the alphavirus genome in place of the capsid-coding region results in the production of a large number of INTRA-coding RNAs and the synthesis of high levels of INTRA in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of INTRA into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA

transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding INTRA.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding INTRA. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs

and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding INTRA. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased INTRA expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding INTRA may be therapeutically useful, and in the treament of disorders associated with decreased INTRA expression or activity, a compound which specifically promotes expression of the polynucleotide encoding INTRA may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding INTRA is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding INTRA are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding INTRA. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5.932,435; Arndt, G.M. et al. (2000) Nucleic Acids

Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such pharmaceutical compositions may consist of INTRA, antibodies to INTRA, and mimetics, agonists, antagonists, or inhibitors of INTRA.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Pharmaceutical compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular delivery of macromolecules comprising INTRA or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, INTRA or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example INTRA or fragments thereof, antibodies of INTRA, and agonists, antagonists or inhibitors of INTRA, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind INTRA may be used for the diagnosis of disorders characterized by expression of INTRA, or in assays to monitor patients being treated with INTRA or agonists, antagonists, or inhibitors of INTRA. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for INTRA include methods which utilize the antibody and a label to detect INTRA in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring INTRA, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of INTRA expression. Normal or standard values for INTRA expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to INTRA under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of INTRA expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding INTRA may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of INTRA may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of INTRA, and to monitor regulation of INTRA levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding INTRA or closely related molecules may be used to identify nucleic acid sequences which encode INTRA. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the

probe identifies only naturally occurring sequences encoding INTRA, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the INTRA encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:53-104 or from genomic sequences including promoters, enhancers, and introns of the INTRA gene.

Means for producing specific hybridization probes for DNAs encoding INTRA include the cloning of polynucleotide sequences encoding INTRA or INTRA derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding INTRA may be used for the diagnosis of disorders associated with expression of INTRA. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, hematopoietic cancer including lymphoma, leukemia, and myeloma; and other cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, adenoma, carcinoma and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal

circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma ; and a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and a hepatic tumor including a nodular hyperplasia,.a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and a gastrointestinal disorder such as esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease, cholecystitis, infections of the intestinal tract, pancreatitis, pancreatic carcinoma, cirrhosis, hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease. The polynucleotide sequences encoding INTRA may be used in Southern or northern analysis, dot blot, or other membrane-based

technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered INTRA expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding INTRA may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding INTRA may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding INTRA in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of INTRA, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding INTRA, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding INTRA may involve the use of PCR. These oligomers may be chemically synthesized, generated

enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding INTRA, or a fragment of a polynucleotide complementary to the polynucleotide encoding INTRA, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding INTRA may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding INTRA are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of INTRA include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be

used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for INTRA, or INTRA or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding INTRA may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man

(OMIM) World Wide Web site. Correlation between the location of the gene encoding INTRA on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, INTRA, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between INTRA and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with INTRA, or fragments thereof, and washed. Bound INTRA is then detected by methods well known in the art. Purified INTRA can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding INTRA specifically compete with a test compound for binding INTRA. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with INTRA.

In additional embodiments, the nucleotide sequences which encode INTRA may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/139,566 (filing date 16 June 1999), U.S. Ser. No. 60/149,640 (filing date 17 August 1999), and U.S. Ser. No. 60/164,417 (filing date 9 November 1999), are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant

plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools,

programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:53-104. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is

much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \times \text{minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding INTRA occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of ABBR Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:8-14 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:8-14 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for

Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

The genetic map locations of SEQ ID NO:8-14 [fill in the specific SEQ ID NOs if not all of the sequences have been mapped] are described in The Invention as ranges, or intervals, of human chromosomes. [Include the following sentence if any of your sequences have more than one map location.] More than one map location is reported for SEQ ID NO:8-14 [fill in specific SEQ ID NO:s], indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:8-14 [fill in specific SEQ ID NO:s] were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Extension of INTRA Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:53-104 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C,

2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:53-104 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:53-104 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a

fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the INTRA-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring INTRA. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of INTRA. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the INTRA-encoding transcript.

X. Expression of INTRA

Expression and purification of INTRA is achieved using bacterial or virus-based expression systems. For expression of INTRA in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory

element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express INTRA upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of INTRA in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding INTRA by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, INTRA is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from INTRA at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified INTRA obtained by these methods can be used directly in the assays shown in Examples XI, XII, and XV.

XI. Demonstration of INTRA Activity

INTRA activity is associated with its ability to form protein-protein complexes and is measured by its ability to regulate growth characteristics of NIH3T3 mouse fibroblast cells. A cDNA encoding INTRA is subcloned into an appropriate eukaryotic expression vector. This vector is transfected into NIH3T3 cells using methods known in the art. Transfected cells are compared with non-transfected cells for the following quantifiable properties: growth in culture to high density, reduced attachment of cells to the substrate, altered cell morphology, and ability to induce tumors when injected into immunodeficient mice. The activity of INTRA is proportional to the extent of increased growth or frequency of altered cell morphology in NIH3T3 cells transfected with INTRA.

Alternatively, INTRA activity is measured by binding of INTRA to radiolabeled formin polypeptides containing the proline-rich region that specifically binds to SH3 containing proteins

(Chan, D.C. et al. (1996) EMBO J. 15: 1045-54). Samples of INTRA are run on SDS-PAGE gels, and transferred onto nitrocellulose by electroblotting. The blots are blocked for 1 hr at room temperature in TBST (137 mM NaCl, 2.7 mM Kcl, 25 mM Tris (pH 8.0) and 0.1% Tween-20) containing non-fat dry milk. Blots are then incubated with TBST containing the radioactive formin polypeptide for 4 hrs to overnight. After washing the blots four times with TBST, the blots are exposed to autoradiographic film. Radioactivity is quantitated by cutting out the radioactive spots and counting them in a radioisotope counter. The amount of radioactivity recovered is proportional to the activity of INTRA in the assay.

Alternatively, INTRA activity is demonstrated by measuring the binding of INTRA to Ca^{2+} using a Ca^{2+} overlay system (Weis, K. et al. (1994) J. Biol. Chem. 269:19142-19150). Purified INTRA is transferred and immobilized onto a nitrocellulose membrane. The membrane is washed three times with buffer (60 mM KCl, 5 mM $MgCl_2$, 10 mM imidazole-HCl, pH 6.8) and incubated in this buffer for 10 minutes with 1 μCi [$^{45}Ca^{2+}$] (NEN-DuPont, Boston, MA). Unbound [$^{45}Ca^{2+}$] is removed from the membrane by washing with water, and the membrane is dried. Membrane-bound [$^{45}Ca^{2+}$] is detected by autoradiography and quantified using image analysis systems and software. INTRA activity is proportional to the amount of [$^{45}Ca^{2+}$] detected on the membrane.

Alternatively, INTRA activity is assayed by measuring the conversion of ^{3}H-cAMP to ^{3}H-adenosine in the presence of INTRA and 5' nucleotidase. INTRA is added to a solution containing 50 mM Tris-HCl pH 7.5, 10 mM $MgCl^2$, 0.1 unit 5'nucleotidase (from Crotalus atrox venom), and 0.0064-2.0 uM ^{3}H- cAMP and the reaction is incubated at 37°C for a time period that would yield less than 15% cAMP hydrolysis in order to avoid non-linearity associated with product inhibition. Soluble radioactivity associated with ^{3}H-adenosine is quantitated using a Beta scintillation counter. The amount of radioactivity recovered is proportional to the activity of INTRA in the reaction.

XII. Functional Assays

INTRA function is assessed by expressing the sequences encoding INTRA at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP;

Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of INTRA on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding INTRA and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding INTRA and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of INTRA Specific Antibodies

INTRA substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the INTRA amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-INTRA activity by, for example, binding the peptide or INTRA to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring INTRA Using Specific Antibodies

Naturally occurring or recombinant INTRA is substantially purified by immunoaffinity chromatography using antibodies specific for INTRA. An immunoaffinity column is constructed by covalently coupling anti-INTRA antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing INTRA are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of INTRA (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/INTRA binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and INTRA is collected.

XV. Identification of Molecules Which Interact with INTRA

INTRA, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled INTRA, washed, and any wells with labeled INTRA complex are assayed. Data obtained using different concentrations of INTRA are used to calculate values for the number, affinity, and association of INTRA with the candidate molecules.

Alternatively, molecules interacting with INTRA are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

INTRA may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	53	129042	TESTNOT01	129042H1 (TESTNOT01), 129042T6 (TESTNOT01), 594163H1 (BRAVUNT02), 1376353T6 (LUNGNOT10), 1968641R6 (BRSTNOT04), 4193335F6 (BRAPDIT01), 5636985H1 (UTRSTMR01)
2	54	778003	COLNNOT05	778003H1 (COLNNOT05), 778003X29 (COLNNOT05), 793138X17 (PROSTUT03), 5533562H1 (HEARFET05)
3	55	1418671	KIDNNOT09	458013F1 (KERANOT01), 461367R6 (KERANOT01), 1418671H1 (KIDNNOT09), 1418671X301D1 (KIDNNOT09), 1452670F1 (PENITUT01), 1455886F1 (COLNFET02), 2921431H1 (SININOT04)
4	56	1456841	COLNFET02	214180X3 (STOMNOT01), 1456841H1 (COLNFET02), 1517021F1 (PANCTUT01), 2280709F6 (COLSUCT01), SBFA01757F1, SBFA04860F1, SBFA03431F1
5	57	2020010	CONNNOT01	520251R1 (MMLR2DT01), 552501H1 (SCORNOT01), 1297508H1 (BRSTNOT07), 1417085H1 (BRAINOT12), 1455946F1 (COLNFET02), 1864670H1 (PROSNOT19), 1922941R6 (BRSTTUT01), 1922941T6 (BRSTTUT01), 1930785H1 (COLNTUT03), 2020010F6 (CONNNOT01), 2020010H1 (CONNNOT01), 2879789H1 (UTRSTUT05), 3324110H1 (PTHYNOT03), 3766286H1 (BRSTNOT24), 4305754H1 (TESTTUT03)
6	58	2149037	BRAINOT09	1382860F1 (BRAITUT08), 1709135F6 (PROSNOT16), 1758155R6 (PITUNOT03), 1861076F6 (PROSNOT19), 2149037H1 (BRAINOT09), 2149037X15F1 (BRAINOT09), 2280366H1 (PROSNON01), 2524642F6 (BRAITUT21), 2590271H1 (LUNGNOT22), 2970418H2 (HEAONOT02), 3084127H1 (BRAIFET01), 4789892T6 (EPIBUNT01)
7	59	2162179	ENDCNOT02	2162179F6 (ENDCNOT02), 2162179H1 (ENDCNOT02), 3865236H1 (BRAITUT07)
8	60	2244706	HIPONON02	2244706H1 (HIPONON02), 3272168F6 (BRAINOT20), SBWA00950V1, SBWA03641V1, SBWA02322V1

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
9	61	2316805	OVARNOT02	363271R6 (PROSNOT01), 855363H1 (NGANNOT01), 1209030T1 (BRSTNOT02), 1265148R1 (SYNORAT05), 1294807F1 (PGANNOT03), 1351585F1 (LATRTUT02), 1852006F6 (LUNGFET03), 2316805H1 (OVARNOT02), 2320867H1 (OVARNOT02), 3563231F6 (SKINNOT05)
10	62	2320010	OVARNOT02	448783H1 (TLYMNOT02), 470134R1 (MMLR1DT01), 720124F1 (SYNOOAT01), 1873477F6 (LEUKNOT02), 2320010H1 (OVARNOT02), 3049510T6 (LUNGNOT25), 3087109F6 (HEAONOT03), 4144881H1 (SINITUT04), 5089346H1 (UTRSTMR01)
11	63	2564901	ADRETUT01	214410F1 (STOMNOT01), 927356R1 (BRAINOT04), 2564901H1 (ADRETUT01)
12	64	2615168	GBLANOT01	1445950F6 (PLACNOT02), 2615168H1 (GBLANOT01), 2746963F6 (LUNGTUT11), 2746963T6 (LUNGTUT11), 3250984H1 (SEMVNOT03), 3459378H1 (293TF1T01), 3831615H1 (PANCNOT17), 4334378H1 (KIDCTMT01), 4818908H1 (PROSTUT17)
13	65	2658329	LUNGTUT09	1210539H1 (BRSTNOT02), 1210539R6 (BRSTNOT02), 1985147R6.comp (LUNGAST01), 2311120R6 (NGANNOT01), 2658329H1 (LUNGTUT09), 2717243F6 (THYRNOT09), 2831384F7 (TLYMNOT03), 3846358H1 (DENDNOT01), 4898171H1 (OVARDIT01)
14	66	2708944	PONSAZT01	309840R6 (TMLR2DT01), 1241166R6 (LUNGNOT03), 1381850H1 (BRAITUT08), 2194624F6 (THYRTUT03), 2212407F6 (SINTFET03), 2708944F6 (PONSAZT01), 2708944H1 (PONSAZT01), 4895659H1 (LIVRTUT12)
15	67	3315012	293TF1T01	532568R6 (BRAINOT03), 1300242F1 (BRSTNOT07), 1329265F1 (PANCNOT07), 1439786H1 (PANCNOT08), 2327916X23C1 (COLNNOT11), 2381037X37C1 (ISLTNOT01), 2381037X39C1 (ISLTNOT01), 3315012H1 (293TF1T01), SAEB00241R1
16	68	4155412	ADRENOT14	555524R6 (SCORNOT01), 4155412F6 (ADRENOT14), 4155412H1 (ADRENOT14), 4943387F6 (BRAIFEN05)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
17	69	4831840	BRAVTXT03	286660H1 (EOSIHET02), 422026H1 (CARCTXT01), 1734445F6 (COLNNOT22), 1734445T6 (COLNNOT22), 1970421F6 (UCMCL5T01), 2512308H1 (CONUTUT01), 4831840H1 (BRAVTXT03)
18	70	5676581	293TF2T01	702633R6 (SYNORAT03), 1000026R1 (BRSTNOT03), 2631308F6 (COLNTUT15), 3012653H1 (MUSCNOT07), 3252744H1 (OVARTUN01), 3315168H2 (293TF2T01), 3530354H1 (BLADNOT09), 4289137H1 (BRABDIR01), 4974749H1 (HELATXT03), 5676581H1 (293TF2T01)
19	71	034159	THP1NOB01	034159H1 (THP1NOB01), 034159X305D3 (THP1NOB01), 406358R6 (EOSIHET02), 1974550F6 (UCMCL5T01), 3471911H1 (LUNGNOT27), 3522363H1 (ESOGTUN01), 4326520H1 (TLYMUNT01), SCJA01020V1, SCIA01764V1
20	72	129023	TESTNOT01	129023R6 (TESTNOT01), 775480R1 (COLNNOT05), 1649938F6 (PROSTUT09), 2518140F6 (BRAITUT21), 2688123H1 (LUNGNOT23), 4306520H1 (MONOTXT01)
21	73	1358940	LUNGNOT09	879273R1 (THYRNOT02), 967670T1 (BRSTNOT05), 1358940F6 (LUNGNOT09), 1358940H1 (LUNGNOT09), 1809259H1 (PROSTUT12), 1818790F6 (PROSNOT20), 1886716F6 (BLADTUT07), 1905126F6 (OVARNOT07), 3508881H1 (CONCNOT01), 3687018F6 (HEAANOT01), 3812474F6 (TONSNOT03)
22	74	1682320	PROSNOT15	1214001T1 (BRSTTUT01), 1259957F1 (MENITUT03), 1375132H1 (LUNGNOT10), 1682320H1 (PROSNOT15), 3137047H1 (SMCCNOT01), 3805984H1 (BLADTUT03), 3806302H1 (BLADTUT03)
23	75	1728263	PROSNOT14	1269315H1 (BRAINOT09), 1453910F1 (PENITUT01), 1728263H1 (PROSNOT14), g2115530
24	76	1867626	SKINBIT01	667711T6 (SCORNOT01), SXYA01116V1, SXYA01833V1, SXYA02442V1

Table 1 (cont.)

25	77	1990126	CORPNOT02	426763T6 (BLADNOT01), 1647316F6 (PROSTUT09), 1757430R6 (PITUNOT03), 1830621F6 (THP1AZT01), 1990126H1 (CORPNOT02), 3250740H1 (SEMVNOT03)
26	78	2104180	BRAITUT02	1350750F1 (LATRTUT02), 1502445F1 (BRAITUT07), 1519125X301D1 (BLADTUT04), 2104180H1 (BRAITUT02), 2733677H1 (OVARTUT04)
27	79	2122241	BRSTNOT07	1402761H1 (LATRTUT02), 1402761T6 (LATRTUT02), 2122241F6 (BRSTNOT07), 2122241H1 (BRSTNOT07), 4989861H1 (LIVRTUT11)
28	80	2580428	KIDNTUT13	157262F1 (THP1PLB02), 1914234X29C1 (PROSTUT04), 1914467X12C1 (PROSTUT04), 1914467X13C1 (PROSTUT04), 1915166X14C1 (PROSTUT04), 2580428H1 (KIDNTUT13), SBKA01222F1
29	81	3397189	UTRSNOT16	759108R6 (BRAITUT02), 1911587T6 (CONNTUT01), 3397189H1 (UTRSNOT16)
30	82	4881249	UTRMTMT01	080470R1 (SYNORAB01), 998242R6 (KIDNTUT01), 4549519H1 (HELAUNT01), 4881249H1 (UTRMTMT01), SXAE01512V1, SXAE02289V1, SXAE00433V1
31	83	431871	EOSINOT03	431871H1 (BRAVUNT02), 460185R1 (KERANOT01), 636514F1 (NEUTGMT01), 1975990T6 (PANCTUT02), 2212046H1 (SINTFET03), 2257310R6 (OVARTUT01), 2300180R6 (BRSTNOT05), 4884920F6 (LUNLTMT01), SCEA00887V1
32	84	526155	EOSINOT02	526155H1 (EOSINOT02), 794168R6 (OVARNOT03), 1260927R1 (SYNORAT05), 1975556F6 (PANCTUT02), 5157385H1 (BRSTTMT02)
33	85	676234	CRBLNOT01	676234H1 (CRBLNOT01), 2241232F6 (PANCTUT02), 2241232T6 (PANCTUT02), 2824092H1 (ADRETUT06), 4248435T6 (BRABDIT01)
34	86	720145	SYNOOAT01	433978H1 (THYRNOT01), 720145H1 (SYNOOAT01), 720145R6 (SYNOOAT01), 2107540T6 (BRAITUT03), 4722278H1 (COLCTUT02)
35	87	1001951	BRSTNOT03	1001951H1 (BRSTNOT03), 1001951R6 (BRSTNOT03), SXYA00708V1, SXYA01879V1, SXYA00520V1, SXYA00731V1, SXYA00926V1
36	88	1243349	LUNGNOT03	050083X316F1 (CHAONOT01), 050083X326F1 (CHAONOT01), 050083X346F1 (CHAONOT01), 050083X350F1 (CHAONOT01), 1243349H1 (LUNGNOT03), 2751089R6 (THP1AZS08), 3773254F6 (BRSTNOT25), 3997530H1 (PROSBPS05), g844357, g1940784, g4539083

Table 1 (cont.)

37	89	1338201	COLNNOT13	256461H1 (HNT2RAT01), 1338201H1 (COLNNOT13), 1338201X12 (COLNNOT13), 1338201X18 (COLNNOT13), 1338201X21 (COLNNOT13), 2078127H1 (ISLTNOT01), g777838, g1146680, g1406379
38	90	1405141	LATRTUT02	189682R6 (CARDNOT01), 551762R6 (SCORNOT01), 1405141X302D1 (LATRTUT02), 1459886X16C1 (COLNFET02), 2601416H1 (UTRSNOT10), 2836108H2 (TLYMNOT03), 3031895F6 (TLYMNOT05), 3127628H1 (LUNGTUT12), 3402733H1 (ESOGNOT03), 4289784F6 (BRABDIR01), 4339406H1 (BRAUNOT02), 4712515H1 (BRAIHCT01), 4746879H2 (SMCRUNT01), 5091792F6 (UTRSTMR01), 5679882H1 (BRAENOT02), 5927661H1 (BRAIFET02)
39	91	1686305	PROSNOT15	499154R6 (NEUTLPT01), 1686305F6 (PROSNOT15), 1686305H1 (PROSNOT15), 2306450R6 (NGANNOT01), 2446232F6 (THP1NOT03), 2446232T6 (THP1NOT03), 3050482H1 (LUNGNOT25), 3694303F6 (LUNGNOT35), 3825239H1 (BRAIHCT01), 3931022H1 (PROSTUT09), 4383527H1 (BRAVUTT02)
40	92	1688972	PROSTUT10	878019H1 (LUNGAST01), 1255436F2 (MENITUT03), 1330287F1 (PANCNOT07), 1400064F6 (BRAITUT08), 1688972H1 (PROSTUT10), 2018742F6 (THP1NOT01), 2047754X12F1 (SININOT01), 3002925H1 (TLYMNOT06), 3744192H1 (THYMNOT08)
41	93	1812494	PROSTUT12	1322590F6 (BLADNOT04), 1684555F6 (PROSNOT15), 2120930H1 (BRSTNOT07), 2266093H1 (UTRSNOT02), 2631470F6 (COLNTUT15), 3980110H1 (LUNGTUT08), 5115462H1 (ENDITXT01), SADA00912R1
42	94	2013853	TESTNOT03	2013853H1 (TESTNOT03), 2013853R6 (TESTNOT03), SXBC01227V1, SCSA04222V1
43	95	2284925	BRAINON01	464655X11 (LATRNOT01), 464655X12 (LATRNOT01), 464655X28 (LATRNOT01), 482019X21 (HNT2RAT01), 1443611R1 (THYRNOT03), 1443611X22 (THYRNOT03), 2284925H1 (BRAINON01), 2882173F6 (UTRSTUT05), 3485205F6 (KIDNNOT31), 3485205T6 (KIDNNOT31), SAAB00144R1

Table 1 (cont.)

44	96	2376728	ISLTNOT01	413593R6 (BRSTNOT01), 823803R1 (PROSNOT06), 860037R1 (BRAITUT03), 1282102F1 (COLNNOT16), 1733518F6 (BRSTTUT08), 2376728F6 (ISLTNOT01), 2376728H1 (ISLTNOT01), 2937285F6 (THYMFET02), 3108296H1 (BRSTTUT15), 3212546H1 (BLADNOT08), 3462704H1 (293TF2T01)
45	97	2790762	COLNTUT16	126628F1 (LUNGNOT01), 126628R1 (LUNGNOT01), 2790762F6 (COLNTUT16), 2790762H1 (COLNTUT16), 4002872H1 (HNT2AZS07), g678705
46	98	2869164	THYRNOT10	1607765F6 (LUNGNOT15), 2869164F6 (THYRNOT10), 2869164H1 (THYRNOT10), 2869164T6 (THYRNOT10), 2890205H1 (LUNGFET04), 2891521F6 (LUNGFET04), 3094580X305D1 (CERVNOT03)
47	99	3317629	PROSBPT03	3166243H1 (SATABT007), 3317629F6 (PROSBPT03), 3421114X302F1 (UCMCNOT04), 4635773F6 (MYEPTXT01), 4635773T6 (MYEPTXT01)
48	100	3870488	BMARNOT03	1670688F6 (BMARNOT03), 3039406T6 (BRSTNOT16), 3870488H1 (BMARNOT03), 4773630H1 (BRAQNOT01)
49	101	3886318	UTRSNOT05	198182F1 (KIDNNOT02), 474711R1 (MMLR1DT01), 733227R1 (LUNGNOT03), 1236870F1 (LUNGFET03), 1502818F1 (BRAITUT07), 3742588H1 (THYMNOT08)
50	102	4043934	LUNGNOT35	4043934F6 (LUNGNOT35), 4043934H1 (LUNGNOT35), g1664159, g2114678, g3665589
51	103	4371445	THYMNOT11	4371445F6 (THYMNOT11), 4371445H1 (THYMNOT11), 4371445T6 (THYMNOT11), g691417
52	104	5527925	KIDNNOT34	878842R1 (THYRNOT02), 1662614F6 (BRSTNOT09), 1820183F6 (GBLATUT01), 2275208H1 (PROSNON01), 2864564H1 (KIDNNOT20), 2890511H1 (LUNGFET04), 4312193H1 (BRAFNOT01), 5175111F6 (EPIBTXT01), 5876074H1 (BRAUNOT01)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences and Motifs	Homologous Sequences	Analytical Methods and Databases
1	446	T24 T144 S251 S384 S404 T114 T118 T121 T172 S181 S247 Y53 Y422	N117 N232	SH3 domain: E387-I441	g2232009, thyroid hormone responsive protein [Rattus norvegicus]. Shah, G.N. et al. (1997) Biochem. J. 327:617-23.	BLAST - GenBank BLAST - DOMO BLIMPS - BLOCKS BLIMPS - PRINTS HMMER - PFAM MOTIFS
2	340	T26 S51 T146 S211 S270 S308 S73 S277 S317 Y71		SH2 domain: W240-Y316	g3738265 SH2 domain-containing protein [Mus musculus]	BLAST - GenBank BLAST - DOMO BLIMPS - PRINTS HMMER - PFAM MOTIFS
3	353	T45 S232 T353 T78 S88 S163 S176 T222 S240 S284 S302 T326 S338 S116 S120 T154 S226 S295 S337		Pleckstrin homology domains: T247-T353 G4-H104 S120-K250	g5381422 pleckstrin 2 [Homo sapiens]	BLAST - GenBank BLAST - PRODOM HMMER - PFAM MOTIFS
4	593	S230 S415 T84 T115 S214 S231 S309 S355 S372 T377 T387 S529 S580 S5 T36 S41 S90 S205 T263 S264 T343 T371 S410 S445 S483 S528 T547	N19 N542	SH3 domain: L453-L507 EPS8 region - SH3/phosphorylation domain: S2-P395	g309217 Eps8 (EGF receptor kinase substrate) [Mus musculus]	BLAST - GenBank BLAST - PRODOM HMMER - PFAM MOTIFS

Table 2

5	358	T42 S82 T204 T233 S261 T271 T279 S285 S330 S55 T102 S153 S254 S353	N338	Ankyrin repeat: G40-G67	g485107 similar to ankyrin repeat region [C. elegans]	BLAST - GenBank HMMER - PFAM MOTIFS
6	749	S137 T401 S406 T407 S580 T29 S140 S148 S149 S287 T336 S342 S360 S511 S551 T627 T29 S104 T368 S480 T616 Y141 Y303	N147 N392 N453 N640	Transmembrane domain: W280-I297 SH3 domain: R483-L537 Probable rabGAP domains: I159-P168 Y200-G205	g1519685 contains similarity to SH3 domains [C. elegans].	BLAST - GenBank BLIMPS - PRINTS BLIMPS - PFAM HMMER - PFAM HMMER MOTIFS
7	139	T51 T113 S106	N31		g169306 calmodulin [Phytophthora infestans]	BLAST - GenBank
8	539	S52 S84 T114 S186 S430 T468 S15 S110 S241 S307 S309 S353 S362 S363 S389 S485 S118 S169 S181 S210 T319 S385 T434 T523 Y208 Y305	N533	Pleckstrin homology domain: R192-A291	g4151807 membrane-associated guanylate kinase-interacting protein 2 (Maguin-2) [Rattus norvegicus]	BLAST - GenBank HMMER - PFAM MOTIFS
9	319	S169 S214 S233 S240 S150	N126	Tumor necrosis factor and nerve growth factor receptors - Conserved domain containing six cysteines: L166-C204	g2809400 Sprouty 2 (antagonist of FGF signaling) [Homo sapiens]	BLAST - GenBank HMMER - PFAM MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences and Motifs	Homologous Sequences	Analytical Methods and Databases
10	747	T194 T344 T561 S655 S45 T58 T60 T74 T81 T171 S287 T294 S446 T526 S608 T610 T733 S126 S133 T165 S170 T190 S234 T251 T429 S470 S492 T522 S546 S735 S741 Y504 Y543	N32 N54 N533 N642		g550420 trg (transcript negatively regulated by thyroid stimulating hormone) [Rattus norvegicus]	
11	266	S62 T76 T183 S222 S4 T5 S256 S260 Y179	N47	Diacylglycerol/phorbol ester binding domain: E177-N223		PROFILESCAN HMMER - PFAM MOTIFS
12	345	T87 S131 S213 T241 S299 S323 T34 T69 T223 S307	N40 N70	Annexin domain: G58-L110 L122-R143 I137-L182 L262-F316 E311-D326 A327-C340	**g3688370, annexin 31 (annexin XXXI) [Homo sapiens]. Morgan, R.O. and Fernandez, M.P. (1998) FEBS Lett. 434:300-304.**	BLAST - GenBank HMMER - PFAM BLIMPS - BLOCKS BLIMPS - PRINTS MOTIFS
13	437	S40 T66 T79 S93 T241 T289 S305 S342 T375 S47 S270 S362 T371 T393			g685183 NGD5 gene product (regulated by opioid treatment) [Murinae gen. sp.]	BLAST - GenBank MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences and Motifs	Homologous Sequences	Analytical Methods and Databases
14	441	S333 S419 T10 T24 T322 S403 S407 S422 T453 S33 S270 S329 T352 S487		Ankyrin repeats: G46-N73 G80-D107	g6460678 ankyrin-related protein [Deinococcus radiodurans].	BLAST - GenBank HMMER - PFAM MOTIFS
15	487	S31 T51 S62 T220 T237 T254 T427 S453 T471 S482 T483 T95 S182	N242 N481	Signal peptide: M1-A28 Histidine acid phosphatase domains: R88-T95 K311-W323 Acid phosphatase-like region: E75-S484	g4105496 multiple inositol polyphosphate phosphatase [Mus musculus].	BLAST - GenBank BLAST - PRODOM BLIMPS - BLOCKS HMMER SPSCAN MOTIFS
16	282	S25 T125 T157 T203 S31 S46 S107 S133 S194 S218 S257	N17 N74 N216		g688297 VDUP1 (1,25-dihydroxy-vitamin D-3 up-regulated polypeptide [Homo sapiens].	BLAST - GenBank MOTIFS
17	581	T147 T327 S477 S41 T119 T123 T129 T209 S232 S243 S257 S299 S341 S347 T366 S371 S142 S220 S223 S237 S276 S323 S399 T472 T487 S518	N221 N358		**g6013191, activating signal cointegrator 1 [H. sapiens]. Kim, H.J.et al. (1999) Mol. Cell. Biol. 19:6323-6332.**	BLAST - GenBank MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences and Motifs	Homologous Sequences	Analytical Methods and Databases
18	530	S23 T46 S219 S221 T267 T268 S290 S303 T370 T382 S406 S446 T2 S31 S195 S339 S358 T375 S379 S399 T424 T445 T504	N43 N99	Signal peptide: M1-S23 WW/rsp5/WWP repeat domain: E123-P153 Trehalase domains: P80-T90 E129-N142	g1255031 FBP 30 (formin binding protein 30) [Mus musculus]	BLAST - GenBank SPSCAN HMMER - PFAM BLIMPS - BLOCKS MOTIFS
19 (034159)	475	S264 T5 T9 S33 S163 S171 S211 S217 S241 T267 S343 S370 T386 S472 S16 S110 S111 S151 S152 S246 T260 S264 T405	N15 N62 N101 N291 N384 N443	Pleckstrin M79-D189 GTPase activator K248-A459	g35013 n-chimaerin	Motifs BLAST_GENBANK HMMER_PFAM BLIMPS_PRINTS BLIMPS_PFAM BLAST_PRODOM BLAST_DOMO
20 (129023)	368	S8 S54 S70 S99 T158 S159 S253 S361 S30 T152 S308	N24 N68 N359	Signal peptide: M1-Q25 WW (signal transduction associated) domain: Y61-P75		Motifs SPSCAN BLIMPS_PRINTS
21 (1358940)	476	S104 S182 T343 S122 T148 T157 T197 S205 T360 S429 T467 T133 T269 T292 T323 S339		EF-hand Calcium binding domain: D231-D421	g3297882 atopy-related autoantigen CALC [H. sapiens].	Motifs BLAST_GENBANK HMMER_PFAM BLAST_PRODOM

Table 2 (cont.)

22 (1682320)	171	T70 T151 S97 Y11 Y24		Leucine zipper: L38-L59 Peptidyl-Prolyl Cis-Trans Isomerase CYP6: L59-F170	g1354207 rof1 FK506 binding protein	Motifs BLAST_GENBANK BLAST_PRODOM BLAST_DOMO
23 (1728263)	163	S16 S39 S56 T101 T112 T131 S148 Y92	N70	EF-hand calcium binding domain: D140-F152	g21209 caltractin **[Scherffelia dubia]**	Motifs BLAST_GENBANK BLAST_PRODOM
24 (1867626)	354	T230 T148 T252 S306 S315 T328 S8 T20 T27 S40 S71 T189 T244 T259 T288	N58 N64 N146 N250	Leucine zipper: L326-L347 ATP-Binding motif: E93-E320 Vasodilator-Stimulated Actin-Binding Phosphoprotein motif: M1-A109	g3834607 homer-1b [Mus musculus]	Motifs BLAST_GENBANK BLAST_PRODOM
25 (1990126)	365	T36 S47 S191 T198 S200 T359 T56 T124 S307 Y80 Y155	N189 N264 N297 N320	Src homology domain 3: R308-L364	g1407657 endophilin II	Motifs BLAST_GENBANK HMMER_PFAM BLIMPS_PRINTS BLAST_DOMO
26 (2104180)	274	T71 S126 T137 S230 S251 T7 S141 S155 Y152	N56	Protein Kinase C2 domain: L55-H135	g3876326 similar to protein kinase C2	Motifs BLAST_GENBANK HMMER_PFAM
27 (2122241)	129	T11 S24 S58 T100 S112 T89		Nascent polypeptide-associated complex alpha chain: G39-T128		Motifs BLAST_DOMO
28 (2580428)	626	S84 S93 S192 S278 T411 S10 S18 T114 S302 S482	N293 N577 N599	Interferon-gamma inducible protein motif: M1-M115, C522-A574	**g4886493 and g6942315, [H. sapiens].**	Motifs BLAST_PRODOM

Table 2 (cont.)

29 (3397189)	157	S7	N97	Signal peptide: M1-S29 Glycosyl hydrolase: L62-L137 Beta D Galactosidase: R28-L153	g2547317 lysosomal beta-galactosidase **WO9914328**	Motifs BLAST_GENBANK SPSCAN HMMER BLIMPS_BLOCKS BLAST_PRODOM
30 (4881249)	383	T7 T26 S90 T62 T81 S102 T363 S3 T210 T256 T286 Y158	N70 N190 N223 N289	WWP (Signal transduction associated proline binding domain):L201-P230	g5059333 ubiquitin ligase	Motifs BLAST_GENBANK HMMER_PFAM BLIMPS_PRINTS
31	478	S186 S202 S270 S354 S455 S9 S94 T175		Signal peptide: M1-A64 Ankyrin repeat: D36-E63 Ankyrin repeat protein domain: Q111-Y174; C285-V447	**g1204166, hypothetical Ank-repeat/BTB-domain protein [Schizosaccharomyces pombe].**	MOTIFS SPSCAN HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM
32	275	S259 T74 T173 S186 T231 S21 T63 T219 S255 S267			COP9 complex subunit 7b [Mus musculus] g3309176	BLAST-Genbank MOTIFS
33	217	T4 T106 S209		Signal peptide: M1-C25 Transmembrane domains: A82-T100; R116-I34 Claudin signature: T21-W30; G49-V55 Q63-L73; D146-V152	claudin-9 protein [Mus musculus] g4325296	BLAST-Genbank MOTIFS SPSCAN HMMER BLIMPS-PRINTS

Table 2 (cont.)

34	74	S6 T58 S54		TPR domain: Y18-P46		MOTIFS HMMER-PFAM BLIMPS-PRODOM
35	367	S309 S24	N240	Transmembrane domain: L257-T277 Armadillo/beta-catenin repeat: 219-252; L252-L265		MOTIFS HMMER BLIMPS-PFAM
36	1113	T17 S43 S609 T755 T52 T215 S239 S287 T307 T313 S504 S510 S535 T536 S635 S688 S804 S812 T856 S863 T884 S938 T983 S996 S1004 S5 T196 S353 S433 T550 S592 S593 S727 T748 S762 S839 T928 S944 T952 T968 S1074 Y23 Y134	N175 N323 N365 N633 N724	PDZ domains: V53-E135; E152-D237 L252-H335; E472-D560 H573-D657; T673-Q754 K989-N1070 SH3 domain repeat: G98-K111 SH3 domain protein signature: V153-G249 GLGF domain: L676-K752	AMPA receptor interacting protein GRIP [Rattus norvegicus] g1890856	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-PFAM BLIMPS-PRODOM BLAST-PRODOM BLAST-DOMO
37	511	S147 S88 S136 T228 T320 S467 T15 T81 T118 T168 S281 S289 S311 S354 S455 T461 T480 T494 Y16 Y114	N86 N116 N315 N316 N355 N403 N425 N429 N478	SH3 domain: Q342-L400	g6563258, insulin receptor tyrosine kinase substrate [Homo sapiens].	BLAST-Genbank MOTIFS HMMER-PFAM

Table 2 (cont.)

38	1177	S421 T936 T96 T121 S164 S209 T256 S277 S325 S374 S388 T397 S435 S443 T456 T519 S662 T669 S727 T901 S983 S1114 S14 T70 S307 S331 S416 S545 T565 S609 T626 T703 S804 S845 S853 S867 T921 S972 T1021 S1108 Y214 Y879 Y171	N84 N1112	Armadillo beta-catenin repeat: I196-L205	trg [Rattus norvegicus] g550420	BLAST-Genbank MOTIFS BLIMPS-PFAM
39	665	S245 T358 S480 T76 S110 S119 S121 T266 S284 S481 S521 S561 S632 S654 S655 S72 S73 S130 T171 S205 T411 S428 T475 S476 T491 S513 S523 T634 Y165 Y567 Y578	N197 N479	TPR domains: L136-P164; Y204-P232 E285-G313; P319-G347 F353-P381 TPR repeat: K137-E252; K286-K395	**g6272680, TPR-containing protein involved in spermatogenesis TPIS [Mus musculus]. Takaishi, M. and Huh, N.H. (1999) Biochem. Biophys. Res. Commun. 264:81-85.**	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-PRODOM BLAST-DOMO
40	125	T119 T67		Signal peptide: M1-A53 SH3 domain: R68-L124 R68-A78; K112-L124		MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS

Table 2 (cont.)

Poly-peptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
41	366	S43 S45 T102 S157 T202 T220 S293 S219 T256 T325 S350 Y237		Signal peptide: M1-S30 Ankyrin repeat: G174-S206	**g289693, homology with isopentenyl-diphosphate-delta-isomerase; [C. elegans]. Sulston, J. et al. (1992) Nature 356:37-41.**	MOTIFS SPSCAN HMMER-PFAM BLIMPS-PFAM
42	173	S16 S42 S48 T67 S100 S111 S152 S86	N126	EF Hands: E22-R53; L57-F85 K94-M122; L135-L163 S-100/IcaBP type calcium binding protein signature: L6-E57; L132-K168 Recoverin family signature: V61-T82; S86-D105 Calmodulin repeat: R25-I79; L119-S157	calcineurin B-like protein (CBLP) [Rattus norvegicus] g220688	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS PROFILESCAN BLAST-DOMO
43	761	S227 S293 S393 S19 S43 T149 T161 S277 T346 T370 T415 T529 T572 S630 T683 S711 T746 S74 S196 S252 S283 S300 T444 T472 T591 S754 Y589	N117 N467 N492 N555	3'5'-cyclic nucleotide phosphodiesterase domain: Y490-H729 D418-W744 3'5'-cyclic nucleotide phosphodiesterase signature: L2-H56; L449-H485 Y490-H501; L516-D556 T572-E610; D657-S711	CAMP-specific cyclic nucleotide phosphodiesterase PDE8 [Mus musculus].	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS PROFILESCAN BLAST-PRODOM BLAST-DOMO

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
44	249	S16 S89 T115 S212 S239 T12 T117 S137 S187 S197 S230 Y208	N84	Pleckstrin homology domain: V35-T131 Rho-GEF domain: L36-C178; E118-D245 FYVE zinc finger: N59-Y64; R171-C183 R202-S212	**g3292902, PUTATIVE RHO/RAC GUANINE NUCLEOTIDE EXCHANGE FACTOR [H. sapiens].**	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-PFAM BLAST-PRODOM
45	247	S109 S44 S53 S123 T138 S167 S95 T98 S127 T220	N90		putative phosphatidyl-inositol 3-kinase [Carassius auratus] g4001815	BLAST-GenBank MOTIFS
46	316	S313 S201 T223 T262 Y186 Y270			g3811347, cytosolic phospholipase A2 beta [Homo sapiens].	BLAST-GenBank MOTIFS
47	334	T119 S97 T182 T244 S316 S317 S324 S60 T72 S97 T179 S187 S290 Y52 Y323	N58 N322	Fes/CIP4 homology domain: G8-L98 SH3 domain/division control protein signature: F6-F287	macrophage actin-associated-tyrosine-phosphorylated protein [Mus musculus] g3947712	BLAST-GenBank MOTIFS HMMER-PFAM BLAST-PRODOM
48	113	T65 S66 T43		SH3 domain: K34-L90	SLP-76 associated protein (TCR-stimulated PK substrate) [Homo sapiens] g2072873	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-PRINTS

Table 2 (cont.)

Poly-peptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
49	264	S18 T76 T163 S181 S167 S223		Wilm's tumor protein signature: D97-P111	SH3 domain binding protein [Rattus norvegicus] g1185397 (P-value= 4.6x10-8).	BLAST-GenBank MOTIFS BLIMPS-PRINTS
50	185	T24 S81 S149 S151 S160 S162 S75 S99 S177 Y176		EF-hands: K101-L129; L143-S171 Recoverin family signature: I23-G42; S93-N112 Calcium binding protein signature: E12-Y104	g1848271, Calcium and integrin binding protein CIB [Homo sapiens]	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM
51	72	T18 S25 T20		Synapse-associated SH3 domain protein signature: M13-E67	homolog of Drosophila discs large protein isoform 1 [Homo sapiens] g558438 (P-value= 7.9x10-9).	BLAST-GenBank MOTIFS BLAST-PRODOM
52	434	S123 T128 S418 S94 T105 S159 S205 T291 S308 S314 T326 T358 S383 S406 S84 T128 T212 Y220	N216 N231	Signal peptide: M1-A50 EF hand: I366-R394 Recoverin family signature: V370-L391	similar to EF hand [C. elegans] g3875264.	BLAST-GenBank MOTIFS SPSCAN HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
53	543-587	Reproductive (0.211) Developmental (0.158) Nervous (0.158)	Cancer (0.421) Cell Proliferation (0.263) Inflammation (0.211)	PBLUESCRIPT
54	273-317 651-695	Nervous (0.462) Gastrointestinal (0.385) Cardiovascular (0.077) Developmental (0.077)	Cancer (0.538) Cell Proliferation (0.308) Inflammation (0.154)	PSPORT1
55	110-154	Developmental (0.174) Gastrointestinal (0.174) Reproductive (0.174)	Cell Proliferation (0.435) Cancer (0.261) Inflammation (0.174)	pINCY
56	273-317 1461-1505	Gastrointestinal (0.821) Reproductive (0.143) Developmental (0.036)	Cancer (0.607) Inflammation (0.286) Cell Proliferation (0.036)	pINCY
57	595-639	Reproductive (0.313) Nervous (0.217) Hematopoietic/Immune (0.120)	Cancer (0.482) Inflammation (0.217) Cell Proliferation (0.169)	pINCY
58	703-747 1297-1341	Reproductive (0.250) Nervous (0.205) Gastrointestinal (0.125)	Cancer (0.509) Cell Proliferation (0.196) Inflammation (0.196)	pINCY
59	417-461	Nervous (0.300) Cardiovascular (0.200) Reproductive (0.200)	Inflammation (0.300) Trauma (0.300) Cancer (0.200) Cell Proliferation (0.200)	pINCY
60	1189-1233	Nervous (1.000)	Neurological (0.500) Trauma (0.333)	PSPORT1
61	272-316	Reproductive (0.314) Gastrointestinal (0.186) Nervous (0.157)	Cancer (0.529) Inflammation (0.200) Cell Proliferation (0.129)	PSPORT1
62	273-317 2055-2099	Hematopoietic/Immune (0.333) Reproductive (0.238) Gastrointestinal (0.167)	Inflammation (0.452) Cancer (0.333) Trauma (0.143)	PSPORT1
63	1-34	Reproductive (0.256) Nervous (0.188) Gastrointestinal (0.120)	Cancer (0.504) Inflammation (0.203) Cell Proliferation (0.195)	PSPORT1

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
64	489-533	Reproductive (0.312) Gastrointestinal (0.125) Nervous (0.125)	Cancer (0.438) Cell Proliferation (0.375) Inflammation (0.188)	pINCY
65	273-317	Reproductive (0.265) Nervous (0.224) Developmental (0.102)	Cancer (0.469) Cell Proliferation (0.286) Inflammation (0.204)	pINCY
66	1028-1072	Cardiovascular (0.286) Nervous (0.200) Reproductive (0.200)	Cancer (0.429) Cell Proliferation (0.171) Inflammation (0.171)	pINCY
67	325-369	Reproductive (0.222) Nervous (0.194) Cardiovascular (0.167) Gastrointestinal (0.167)	Cancer (0.472) Cell Proliferation (0.333) Inflammation (0.139)	pINCY
68	921-965	Endocrine (0.250) Musculoskeletal (0.250) Reproductive (0.250) Urologic (0.250)	Cancer (0.750) Trauma (0.250)	pINCY
69	1029-1073	Reproductive (0.216) Gastrointestinal (0.176) Hematopoietic/Immune (0.157)	Cancer (0.510) Inflammation (0.275) Cell Proliferation (0.118)	pINCY
70	1405-1449	Hematopoietic/Immune (0.200) Nervous (0.200) Gastrointestinal (0.160) Reproductive (0.160)	Cancer (0.360) Inflammation (0.360) Cell Proliferation (0.200)	pINCY
71	280-324	Hematopoietic/Immune (0.500) Gastrointestinal (0.092) Reproductive (0.092)	Cancer (0.364) Inflammation (0.295) Cell proliferation (0.205)	pBLUESCRIPT
72	380-424	Reproductive (0.227) Gastrointestinal (0.205) Cardiovascular (0.114)	Cancer (0.455) Inflammation (0.364) Trauma (0.045)	pBLUESCRIPT

Table 3 (cont.)

73	433-477	Nervous (0.241) Reproductive (0.231) Gastrointestinal (0.130)	Cancer (0.398) Inflammation (0.333)	pINCY
74	786-830	Reproductive (0.342) Nervous (0.210)	Cancer (0.474) Cell proliferation (0.184) Inflammation (0.105)	pINCY
75	1-47	Gastrointestinal (0.286) Reproductive (0.286) Developmental (0.143) Hematopoietic/Immune (0.143)	Cancer (0.571) Cell proliferation (0.286) Inflammation (0.143)	pINCY
76	380-424	Nervous (0.300) Reproductive (0.200)	Inflammation (0.400) Cancer (0.200) Cell proliferation (0.200)	pINCY
77	30-74	Gastrointestinal (0.222) Reproductive (0.222) Cardiovascular (0.153) Nervous (0.153)	Inflammation (0.375) Cancer (0.361) Cell proliferation (0.139)	pINCY
78	487-531	Nervous (0.300) Reproductive (0.183) Cardiovascular (0.117)	Cancer (0.433) Inflammation (0.200) Neurological (0.133)	pSPORT1
79	595-639	Reproductive (0.305) Nervous (0.179) Gastrointestinal (0.126)	Cancer (0.526) Inflammation (0.326) Cell proliferation (0.179)	pINCY
80	109-153	Reproductive (0.235) Hematopoietic/Immune (0.216) Nervous (0.157)	Cancer (0.529) Inflammation (0.255)	pINCY
81	109-153	Gastrointestinal (0.286) Musculoskeletal (0.286) Reproductive (0.286)	Cancer (0.571) Inflammation (0.286)	pINCY
82	163-207	Reproductive (0.424) Gastrointestinal (0.152) Nervous (0.121)	Cancer (0.424) Inflammation (0.242) Cell proliferation (0.182)	pINCY
83	496-540	Reproductive (0.242) Nervous (0.182) Hematopoietic/Immune (0.167)	Cancer (0.455) Inflammation/Trauma (0.364) Cell Proliferation (0.152)	PSPORT1
84	1022-1066	Reproductive (0.248) Nervous (0.208) Cardiovascular (0.136)	Cancer (0.464) Inflammation/Trauma (0.304) Cell Proliferation (0.184)	PSPORT1

Table 3 (cont.)

85	39-83	Nervous (0.286) Endocrine (0.143) Gastrointestinal (0.143) Hematopoietic/Immune (0.143) Reproductive (0.143)	Cancer (0.571) Inflammation/Trauma (0.286) Neurological (0.143)	PSPORT1
86	471-515	Hematopoietic/Immune (0.167) Musculoskeletal (0.167) Reproductive (0.167)	Cancer (0.556) Cell Proliferation (0.167) Inflammation/Trauma (0.167)	PSPORT1
87	595-639 982-1026	Reproductive (0.294) Cardiovascular (0.176) Gastrointestinal (0.176)	Cancer (0.706) Inflammation/Trauma (0.294) Cell Proliferation (0.118)	PSPORT1
88	1101-1163	Reproductive (0.625) Gastrointestinal (0.250) Cardiovascular (0.125)	Cancer (0.750) Inflammation/Trauma (0.250)	PSPORT1
89	1245-1289	Gastrointestinal (0.387) Reproductive (0.355) Cardiovascular (0.065)	Cancer (0.548) Inflammation/Trauma (0.323) Cell Proliferation (0.161)	pINCY
90	3720-3764	Nervous (0.328) Gastrointestinal (0.121) Reproductive (0.121)	Cancer (0.397) Inflammation/Trauma (0.310) Cell Proliferation (0.155)	pINCY
91	659-703 1622-1666	Hematopoietic/Immune (0.273) Nervous (0.182) Cardiovascular (0.121) Reproductive (0.121)	Cancer (0.455) Cell Proliferation (0.333) Inflammation/Trauma (0.303)	pINCY
92	104-148	Reproductive (0.310) Nervous (0.241) Developmental (0.138) Gastrointestinal (0.138)	Cancer (0.483) Inflammation/Trauma (0.241) Cell Proliferation (0.172)	pINCY
93	820-864	Reproductive (0.340) Cardiovascular (0.120) Nervous (0.120)	Inflammation/Trauma (0.440) Cancer (0.400) Cell Proliferation (0.160)	pINCY
94	504-554	Reproductive (1.000)	Inflammation/Trauma (1.000)	PBLUESCRIPT
95	198-242	Reproductive (0.424) Nervous (0.273)	Cancer (0.576) Inflammation/Trauma (0.182)	PSPORT1
96	307-351 712-756	Reproductive (0.412) Hematopoietic/Immune (0.137) Cardiovascular (0.118) Gastrointestinal (0.118)	Cancer (0.608) Inflammation/Trauma (0.275) Cell Proliferation (0.098)	pINCY

Table 3 (cont.)

97	433-477	Developmental (0.200) Reproductive (0.200) Cardiovascular (0.133) Gastrointestinal (0.133) Nervous (0.133)	Cell Proliferation (0.400) Cancer (0.333) Inflammation/Trauma (0.200)	pINCY
98	474-1018	Cardiovascular (0.190) Reproductive (0.190) Hematopoietic/Immune (0.143) Musculoskeletal (0.143)	Cancer (0.381) Inflammation/Trauma (0.333)	pINCY
99	422-466 998-1042	Hematopoietic/Immune (0.667) Reproductive (0.222) Developmental (0.111)	Inflammation/Trauma (0.556) Cancer (0.222) Cell Proliferation (0.222)	pINCY
100	444-488	Hematopoietic/Immune (0.455) Nervous (0.182) Cardiovascular (0.091)	Inflammation/Trauma (0.546) Cancer (0.182) Cell Proliferation (0.182)	pINCY
101	1578-1622	Reproductive (0.250) Nervous (0.170) Gastrointestinal (0.156)	Cancer (0.482) Inflammation/Trauma (0.345) Cell Proliferation (0.167)	pINCY
102	15-59	Cardiovascular (1.000)	Cancer (1.000)	pINCY
103	487-531	Hematopoietic/Immune (1.000)		pINCY
104	967-1011	Reproductive (0.235) Nervous (0.191) Gastrointestinal (0.147)	Cancer (0.515) Inflammation/Trauma (0.294) Cell Proliferation (0.118)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
53	TESTNOT01	The library was constructed using RNA isolated from the testicular tissue of a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
54	COLNNOT05	The library was constructed using RNA isolated from the sigmoid colon tissue of a 40-year-old Caucasian male during a partial colectomy. Pathology indicated Crohn's disease involving the proximal colon and including the cecum. The ascending and transverse colon displayed linear ulcerations and skip lesions. Transmural inflammation was present.
55	KIDNNOT09	The library was constructed using RNA isolated from the kidney tissue of a Caucasian male fetus who died at 23 weeks' gestation.
56	COLNFET02	The library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus who died at 20 weeks' gestation.
57	CONNNOT01	The library was constructed using RNA isolated from mesentery fat tissue obtained from a 71-year-old Caucasian male during a partial colectomy and permanent colostomy. Family history included atherosclerotic coronary artery disease, myocardial infarction, and extrinsic asthma.
58	BRAINOT09	The library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who died at 23 weeks' gestation.
59	ENDCNOT02	The library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a 30-year-old Caucasian female.
60	HIPONON02	This normalized library was constructed using 1.13 million independent clones from a hippocampus library. RNA was isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
61	OVARNOT02	The library was constructed using RNA isolated from ovarian tissue removed from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, myocardial infarction, hypercholesterolemia, hypotension, and arthritis.
62	OVARNOT02	The library was constructed using RNA isolated from ovarian tissue removed from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, myocardial infarction, hypercholesterolemia, hypotension, and arthritis.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
63	ADRETUT01	The library was constructed using RNA isolated from right adrenal tumor tissue removed from a 50-year-old Turkish male during a unilateral adrenalectomy. Pathology indicated a metastatic renal cell carcinoma that formed a circumscribed, spongy, hemorrhagic nodule situated in the region of the medulla. The patient presented with corticoadrenal insufficiency, incisional hernia, and non-alcoholic steato hepatitis. Patient history included renal cell carcinoma. Family history included liver cancer.
64	GBLANOT01	The library was constructed using RNA isolated from diseased gallbladder tissue removed from a 53-year-old Caucasian female during a cholecystectomy. Pathology indicated mild chronic cholecystitis and cholelithiasis with approximately 150 mixed gallstones. Family history included benign hypertension.
65	LUNGTUT09	The library was constructed using RNA isolated from lung tumor tissue removed from a 68-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma and a metastatic tumor. Patient history included type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use.
66	PONSAZT01	The library was constructed using RNA isolated from diseased pons tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
67	293TF1T01	The library was constructed using RNA isolated from a transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were transformed with adenovirus 5 DNA.
68	ADRENOT14	The library was constructed using RNA isolated from adrenal gland tissue removed from an 8-year-old Black male who died from anoxia.
69	BRAVTXT03	The library was constructed using RNA isolated from treated astrocytes removed from the brain of a female fetus who died at 22 weeks' gestation. The cells were treated with tumor necrosis factor (TNF) alpha and interleukin 1 (IL-1), 10ng/ml each for 24 hours.

Table 4 (cont.)

70	293TF2T01	The library was constructed using RNA isolated from a treated, transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were treated with 5-aza-2'-deoxycytidine and transformed with adenovirus 5 DNA.
71	THP1NOB01	Library was constructed using RNA isolated from cultured, unstimulated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia. RNA was isolated from 2x108 cells using GuSCN lysis, followed by DNAse treatment.
72	TESTNOT01	Library was constructed using RNA isolated from the testicular tissue of a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
73	LUNGNOT09	Library was constructed using RNA isolated from the lung tissue of a 23-week-old Caucasian male fetus. The pregnancy was terminated following a diagnosis by ultrasound of infantile polycystic kidney disease.
74	PROSNOT15	Library was constructed using RNA isolated from diseased prostate tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 2+3). The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer, secondary bone cancer, and benign hypertension.
75	PROSNOT14	Library was constructed using RNA isolated from diseased prostate tissue removed from a 60-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+4). Family history included benign hypertension, cerebrovascular disease, and arteriosclerotic coronary artery disease.
76	SKINBIT01	Library was constructed using RNA isolated from diseased skin tissue of the left lower leg. Patient history included erythema nodosum of the left lower leg.
77	CORPNOT02	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
78	BRAITUT02	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.

Table 4 (cont.)

79	BRSTNOT07	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
80	KIDNTUT13	Library was constructed using RNA isolated from kidney tumor tissue removed from a 51-year-old Caucasian female during a nephroureterectomy. Pathology indicated a grade 3 renal cell carcinoma. Family history included calculus of the kidney, colon cancer, and type II diabetes.
81	UTRSNOT16	Library was constructed using RNA isolated from uterine endometrial tissue removed from a 48-year-old Caucasian female during a vaginal hysterectomy, rectocele repair, and bilateral salpingo-oophorectomy. Pathology indicated chronic cervicitis, and the endometrium was weakly proliferative. Pathology for the associated tumor tissue indicated a single submucosal leiomyoma. Patient history included hyperlipidemia and meningitis. Family history included benign hypertension, hyperlipidemia, atrial fibrillation, atherosclerotic coronary artery disease, and type II diabetes.
82	UTRMTMT01	Library was constructed using RNA isolated from myometrial tissue removed from a 45-year-old Caucasian female during vaginal hysterectomy and bilateral salpingo-oophorectomy. Pathology indicated the myometrium was negative for tumor. Pathology for the associated tumor tissue indicated multiple (23) subserosal, intramural, and submucosal leiomyomata. Patient history included extrinsic asthma without status asthmaticus and normal delivery. Family history included cerebrovascular disease, depression, and atherosclerotic coronary artery disease.
83	EOSINOT03	This library was constructed using RNA isolated from pooled diseased eosinophils obtained from allergic asthmatic individuals.
84	EOSINOT02	This library was constructed using RNA isolated from pooled eosinophils obtained from allergic asthmatic individuals.
85	CRBLNOT01	This library was constructed using RNA isolated from the cerebellum tissue of a 69-year-old Caucasian male who died from chronic obstructive pulmonary disease. Patient history included myocardial infarction, hypertension, and osteoarthritis.
86	SYNOOAT01	This library was constructed using RNA isolated from the knee synovial membrane tissue of an 82-year-old female with osteoarthritis.
87	BRSTNOT03	This library was constructed using RNA isolated from diseased breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Patient history included kidney infection and condyloma acuminatum.

		Family history included benign hypertension, hyperlipidemia, and a malignant neoplasm of the colon.
88	LUNGNOT03	This library was constructed using RNA isolated from lung tissue of a 79-year-old Caucasian male. Pathology for the associated tumor tissue indicated grade 4 carcinoma. Patient history included a benign prostate neoplasm and atherosclerosis.
89	COLNNOT13	This library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
90	LATRTUT02	This library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
91	PROSNOT15	This library was constructed using RNA isolated from diseased prostate tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 2+3). The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer, secondary bone cancer, and benign hypertension.
92	PROSTUT10	This library was constructed using RNA isolated from prostatic tumor tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 2+3). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer and secondary bone cancer.
93	PROSTUT12	This library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
94	TESTNOT03	This library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
95	BRAINON01	This library was constructed and normalized from 4.88 million independent clones from a brain library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain.

Table 4 (cont.)

96	ISLTNOT01	This library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.
97	COLNTUT16	This library was constructed using RNA isolated from colon tumor tissue obtained from a 60-year-old Caucasian male during a left hemicolectomy. Pathology indicated an invasive grade 2 adenocarcinoma, forming a sessile mass. Patient history included thrombophlebitis, inflammatory polyarthropathy, prostatic inflammatory disease, and depressive disorder. Previous surgeries included resection of the rectum. Family history included atherosclerotic coronary artery disease and colon cancer.
98	THYRNOT10	This library was constructed using RNA isolated from the diseased left thyroid tissue removed from a 30-year-old Caucasian female during a unilateral thyroid lobectomy and parathyroid reimplantation. Pathology indicated lymphocytic thyroiditis.
99	PROSBPT03	This library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy and regional lymph node excision. Pathology indicated benign prostatic hyperplasia (BPH). Pathology for the associated tumor indicated adenocarcinoma, Gleason grade 3+3. The patient presented with elevated prostate specific antigen (PSA), benign hypertension, and hyperlipidemia. Family history included cerebrovascular disease, benign hypertension and prostate cancer.
100	BMARNOT03	This library was constructed using RNA isolated from the left tibial bone marrow tissue of a 16-year-old Caucasian male during a partial left tibial ostectomy with free skin graft. Patient history included an abnormality of the red blood cells. Family history included osteoarthritis.
101	UTRSNOT05	This library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.

Table 4 (cont.)

102	LUNGNOT35	This library was constructed using RNA isolated from lung tissue removed from a 62-year-old Caucasian female. Pathology for the associated tumor tissue indicated a grade 1 spindle cell carcinoid forming a nodule. Patient history included depression, thrombophlebitis, and hyperlipidemia. Family history included cerebrovascular disease, atherosclerotic coronary artery disease, breast cancer, colon cancer,type II diabetes, and malignant skin melanoma.
103	THYMNOT11	This library was constructed using RNA isolated from thymus tissue removed from a 2-year-old Caucasian female during a thymectomy and patch closure of left atrioventricular fistula. The patient presented with congenital heart abnormalities. Patient history included double inlet left ventricle and a rudimentary right ventricle, pulmonary hypertension, cyanosis, subaortic stenosis, seizures, and a fracture of the skull base. Family history included reflux neuropathy.
104	KIDNNOT34	This library was constructed using RNA isolated from left kidney tissue obtained from an 8-year-old Caucasian male who died from an intracranial hemorrhage.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:96, SEQ ID

NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, and SEQ ID NO:104.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104,

b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:53-104,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe

specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52.

18. A method for treating a disease or condition associated with decreased expression of functional INTRA, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A pharmaceutical composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional INTRA, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional INTRA, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.
YUE, Henry
TANG, Y. Tom
HILLMAN, Jennifer L.
LAL, Preeti
BANDMAN, Olga
BAUGHN, Mariah R.
AZIMZAI, Yalda
YANG, Junming
REDDY, Roopa
LU, Dyung Aina M.

<120> INTRACELLULAR SIGNALING MOLECULES

<130> PF-0733 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/149,640; 60/164,417
<151> 1999-08-17; 1999-11-09

<160> 104

<170> PERL Program

<210> 1
<211> 446
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 129042CD1

<400> 1
```
Met Ala Asp Val Gln Met Leu Leu Glu Glu Glu Ile Pro Gly Gly
 1               5                  10                  15
Arg Arg Ala Leu Phe Asp Ser Tyr Thr Asn Leu Glu Arg Val Ala
                20                  25                  30
Asp Tyr Cys Glu Asn Asn Tyr Ile Gln Ser Ala Asp Lys Gln Arg
                35                  40                  45
Ala Leu Glu Glu Thr Lys Ala Tyr Thr Thr Gln Ser Leu Ala Ser
                50                  55                  60
Val Ala Tyr Leu Ile Asn Thr Leu Ala Asn Asn Val Leu Gln Met
                65                  70                  75
Leu Asp Ile Gln Ala Ser Gln Leu Arg Arg Met Glu Ser Ser Ile
                80                  85                  90
Asn His Ile Ser Gln Thr Val Asp Ile His Lys Glu Lys Val Ala
                95                 100                 105
Arg Arg Glu Ile Gly Ile Leu Thr Thr Asn Lys Asn Thr Ser Arg
               110                 115                 120
Thr His Lys Ile Ile Ala Pro Ala Asn Leu Glu Arg Pro Val Arg
               125                 130                 135
Tyr Ile Arg Lys Pro Ile Asp Tyr Thr Ile Leu Asp Asp Ile Gly
               140                 145                 150
His Gly Val Lys Val Ser Thr Gln Asn Met Lys Met Gly Gly Leu
               155                 160                 165
Pro Arg Thr Thr Pro Pro Thr Gln Lys Pro Pro Ser Pro Pro Met
               170                 175                 180
Ser Gly Lys Gly Thr Leu Gly Arg His Ser Pro Tyr Arg Thr Leu
               185                 190                 195
Glu Pro Val Arg Pro Pro Val Val Pro Asn Asp Tyr Val Pro Ser
               200                 205                 210
```

```
Pro Thr Arg Asn Met Ala Pro Ser Gln Gln Ser Pro Val Arg Thr
                215                 220                 225
Ala Ser Val Asn Gln Arg Asn Arg Thr Tyr Ser Ser Ser Gly Ser
                230                 235                 240
Ser Gly Gly Ser His Pro Ser Ser Arg Ser Ser Ser Arg Glu Asn
                245                 250                 255
Ser Gly Ser Gly Ser Val Gly Val Pro Ile Ala Val Pro Thr Pro
                260                 265                 270
Ser Pro Pro Ser Val Phe Pro Gly His Pro Val Gln Phe Tyr Ser
                275                 280                 285
Met Asn Arg Pro Ala Ser Arg His Thr Pro Pro Thr Ile Gly Gly
                290                 295                 300
Ser Leu Pro Tyr Arg Arg Pro Pro Ser Ile Thr Ser Gln Thr Ser
                305                 310                 315
Leu Gln Asn Gln Met Asn Gly Gly Pro Phe Tyr Ser Gln Asn Pro
                320                 325                 330
Val Ser Asp Thr Pro Pro Pro Pro Pro Pro Val Glu Glu Pro Val
                335                 340                 345
Phe Asp Glu Ser Pro Pro Pro Pro Pro Pro Pro Glu Asp Tyr Glu
                350                 355                 360
Glu Glu Glu Ala Ala Val Val Glu Tyr Ser Asp Pro Tyr Ala Glu
                365                 370                 375
Glu Asp Pro Pro Trp Ala Pro Arg Ser Tyr Leu Glu Lys Val Val
                380                 385                 390
Ala Ile Tyr Asp Tyr Thr Lys Asp Lys Glu Asp Glu Leu Ser Phe
                395                 400                 405
Gln Glu Gly Ala Ile Ile Tyr Val Ile Lys Lys Asn Asp Asp Gly
                410                 415                 420
Trp Tyr Glu Gly Val Met Asn Gly Val Thr Gly Leu Phe Pro Gly
                425                 430                 435
Asn Tyr Val Glu Ser Ile Met His Tyr Ser Glu
                440                 445
```

<210> 2
<211> 340
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 778003CD1

<400> 2

```
Met Ala Lys Trp Leu Arg Asp Tyr Leu Ser Phe Gly Gly Arg Arg
 1              5                   10                  15
Pro Pro Pro Gln Pro Pro Thr Pro Asp Tyr Thr Glu Ser Asp Ile
                20                  25                  30
Leu Arg Ala Tyr Arg Ala Gln Lys Asn Leu Asp Phe Glu Asp Pro
                35                  40                  45
Tyr Glu Asp Ala Glu Ser Arg Leu Glu Pro Asp Pro Ala Gly Pro
                50                  55                  60
Gly Asp Ser Lys Asn Pro Gly Asp Ala Lys Tyr Gly Ser Pro Lys
                65                  70                  75
His Arg Leu Ile Lys Val Glu Ala Ala Asp Met Ala Arg Ala Lys
                80                  85                  90
Ala Leu Leu Gly Gly Pro Gly Glu Glu Leu Glu Ala Asp Thr Glu
                95                  100                 105
Tyr Leu Asp Pro Phe Asp Ala Gln Pro His Pro Ala Pro Pro Asp
                110                 115                 120
Asp Gly Tyr Met Glu Pro Tyr Asp Ala Gln Trp Val Met Ser Glu
                125                 130                 135
Leu Pro Gly Arg Gly Val Gln Leu Tyr Asp Thr Pro Tyr Glu Glu
                140                 145                 150
Gln Asp Pro Glu Thr Ala Asp Gly Pro Pro Ser Gly Gln Lys Pro
                155                 160                 165
Arg Gln Ser Arg Met Pro Gln Glu Asp Glu Arg Pro Ala Asp Glu
                170                 175                 180
Tyr Asp Gln Pro Trp Glu Trp Lys Lys Asp His Ile Ser Arg Ala
```

185 190 195
Phe Ala Val Gln Phe Asp Ser Pro Glu Trp Glu Arg Thr Pro Gly
200 205 210
Ser Ala Lys Glu Leu Arg Arg Pro Pro Pro Arg Ser Pro Gln Pro
215 220 225
Ala Glu Arg Val Asp Pro Ala Leu Pro Leu Glu Lys Gln Pro Trp
230 235 240
Phe His Gly Pro Leu Asn Arg Ala Asp Ala Glu Ser Leu Leu Ser
245 250 255
Leu Cys Lys Glu Gly Ser Tyr Leu Val Arg Leu Ser Glu Thr Ser
260 265 270
Pro Gln Asp Cys Ser Leu Ser Leu Arg Ser Ser Gln Gly Phe Leu
275 280 285
His Leu Lys Phe Ala Arg Thr Arg Glu Asn Gln Val Val Leu Gly
290 295 300
Gln His Ser Gly Pro Phe Pro Ser Val Pro Glu Leu Val Leu His
305 310 315
Tyr Ser Ser Arg Pro Leu Pro Val Gln Gly Ala Glu His Leu Ala
320 325 330
Leu Leu Tyr Pro Val Val Thr Gln Thr Pro
335 340

<210> 3
<211> 353
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1418671CD1

<400> 3
Met Glu Asp Gly Val Leu Lys Glu Gly Phe Leu Val Lys Arg Gly
1 5 10 15
His Ile Val His Asn Trp Lys Ala Arg Trp Phe Ile Leu Arg Gln
20 25 30
Asn Thr Leu Val Tyr Tyr Lys Leu Glu Gly Gly Arg Arg Val Thr
35 40 45
Pro Pro Lys Gly Arg Ile Leu Leu Asp Gly Cys Thr Ile Thr Cys
50 55 60
Pro Cys Leu Glu Tyr Glu Asn Arg Pro Leu Leu Ile Lys Leu Lys
65 70 75
Thr Gln Thr Ser Thr Glu Tyr Phe Leu Glu Ala Cys Ser Arg Glu
80 85 90
Glu Arg Asp Ala Trp Ala Phe Glu Ile Thr Gly Ala Ile His Ala
95 100 105
Gly Gln Pro Gly Lys Val Gln Gln Leu His Ser Leu Arg Asn Ser
110 115 120
Phe Lys Leu Pro Pro His Ile Ser Leu His Arg Ile Val Asp Lys
125 130 135
Met His Asp Ser Asn Thr Gly Ile Arg Ser Ser Pro Asn Met Glu
140 145 150
Gln Gly Ser Thr Tyr Lys Lys Thr Phe Leu Gly Ser Ser Leu Val
155 160 165
Asp Trp Leu Ile Ser Asn Ser Phe Thr Ala Ser Arg Leu Glu Ala
170 175 180
Val Thr Leu Ala Ser Met Leu Met Glu Glu Asn Phe Leu Arg Pro
185 190 195
Val Gly Val Arg Ser Met Gly Ala Ile Arg Ser Gly Asp Leu Ala
200 205 210
Glu Gln Phe Leu Asp Asp Ser Thr Ala Leu Tyr Thr Phe Ala Glu
215 220 225
Ser Tyr Lys Lys Lys Ile Ser Pro Lys Glu Glu Ile Ser Leu Ser
230 235 240
Thr Val Glu Leu Ser Gly Thr Val Val Lys Gln Gly Tyr Leu Ala
245 250 255
Lys Gln Gly His Lys Arg Lys Asn Trp Lys Val Arg Arg Phe Val
260 265 270

Leu Arg Lys Asp Pro Ala Phe Leu His Tyr Tyr Asp Pro Ser Lys
275 280 285
Glu Glu Asn Arg Pro Val Gly Gly Phe Ser Leu Arg Gly Ser Leu
290 295 300
Val Ser Ala Leu Glu Asp Asn Gly Val Pro Thr Gly Val Lys Gly
305 310 315
Asn Val Gln Gly Asn Leu Phe Lys Val Ile Thr Lys Asp Asp Thr
320 325 330
His Tyr Tyr Ile Gln Ala Ser Ser Lys Ala Glu Arg Ala Glu Trp
335 340 345
Ile Glu Ala Ile Lys Lys Leu Thr
350

<210> 4
<211> 593
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1456841CD1

<400> 4
Met Ser Arg Pro Ser Ser Arg Ala Ile Tyr Leu His Arg Lys Glu
1 5 10 15
Tyr Ser Gln Asn Leu Thr Ser Glu Pro Thr Leu Leu Gln His Arg
20 25 30
Val Glu His Leu Met Thr Cys Lys Gln Gly Ser Gln Arg Val Gln
35 40 45
Gly Pro Glu Asp Ala Leu Gln Lys Leu Phe Glu Met Asp Ala Gln
50 55 60
Gly Arg Val Trp Ser Gln Asp Leu Ile Leu Gln Val Arg Asp Gly
65 70 75
Trp Leu Gln Leu Leu Asp Ile Glu Thr Lys Glu Glu Leu Asp Ser
80 85 90
Tyr Arg Leu Asp Ser Ile Gln Ala Met Asn Val Ala Leu Asn Thr
95 100 105
Cys Ser Tyr Asn Ser Ile Leu Ser Ile Thr Val Gln Glu Pro Gly
110 115 120
Leu Pro Gly Thr Ser Thr Leu Leu Phe Gln Cys Gln Glu Val Gly
125 130 135
Ala Glu Arg Leu Lys Thr Ser Leu Gln Lys Ala Leu Glu Glu Glu
140 145 150
Leu Glu Gln Arg Pro Arg Leu Gly Gly Leu Gln Pro Ser Gln Asp
155 160 165
Arg Trp Arg Gly Pro Ala Met Glu Arg Pro Leu Pro Met Glu Gln
170 175 180
Ala Arg Tyr Leu Glu Pro Gly Ile Pro Pro Glu Gln Pro His Gln
185 190 195
Arg Thr Leu Glu His Ser Leu Pro Pro Ser Pro Arg Pro Leu Pro
200 205 210
Arg His Thr Ser Ala Arg Glu Pro Ser Ala Phe Thr Leu Pro Pro
215 220 225
Pro Arg Arg Ser Ser Ser Pro Glu Asp Pro Glu Arg Asp Glu Glu
230 235 240
Val Leu Asn His Val Leu Arg Asp Ile Glu Leu Phe Met Gly Lys
245 250 255
Leu Glu Lys Ala Gln Ala Lys Thr Ser Arg Lys Lys Lys Phe Gly
260 265 270
Lys Lys Asn Lys Asp Gln Gly Gly Leu Thr Gln Ala Gln Tyr Ile
275 280 285
Asp Cys Phe Gln Lys Ile Lys Tyr Ser Phe Asn Leu Leu Gly Arg
290 295 300
Leu Ala Thr Trp Leu Lys Glu Thr Ser Ala Pro Glu Leu Val His
305 310 315
Ile Leu Phe Lys Ser Leu Asn Phe Ile Leu Ala Arg Cys Pro Glu
320 325 330
Ala Gly Leu Ala Ala Gln Val Ile Ser Pro Leu Leu Thr Pro Lys

```
                    335                     340                     345
Ala Ile Asn Leu Leu Gln Ser Cys Leu Ser Pro Pro Glu Ser Asn
                    350                     355                     360
Leu Trp Met Gly Leu Gly Pro Ala Trp Thr Thr Ser Arg Ala Asp
                    365                     370                     375
Trp Thr Gly Asp Glu Pro Leu Pro Tyr Gln Pro Thr Phe Ser Asp
                    380                     385                     390
Asp Trp Gln Leu Pro Glu Pro Ser Ser Gln Ala Pro Leu Gly Tyr
                    395                     400                     405
Gln Asp Pro Val Ser Leu Arg Arg Gly Ser His Arg Leu Gly Ser
                    410                     415                     420
Thr Ser His Phe Pro Gln Glu Lys Thr His Asn His Asp Pro Gln
                    425                     430                     435
Pro Gly Asp Pro Asn Ser Arg Pro Ser Ser Pro Lys Pro Ala Gln
                    440                     445                     450
Pro Ala Leu Lys Met Gln Val Leu Tyr Glu Phe Glu Ala Arg Asn
                    455                     460                     465
Pro Arg Glu Leu Thr Val Val Gln Gly Glu Lys Leu Glu Val Leu
                    470                     475                     480
Asp His Ser Lys Arg Trp Trp Leu Val Lys Asn Glu Ala Gly Arg
                    485                     490                     495
Ser Gly Tyr Ile Pro Ser Asn Ile Leu Glu Pro Leu Gln Pro Gly
                    500                     505                     510
Thr Pro Gly Thr Gln Gly Gln Ser Pro Ser Arg Val Pro Met Leu
                    515                     520                     525
Arg Leu Ser Ser Arg Pro Glu Glu Val Thr Asp Trp Leu Gln Ala
                    530                     535                     540
Glu Asn Phe Ser Thr Ala Thr Val Arg Thr Leu Gly Ser Leu Thr
                    545                     550                     555
Gly Ser Gln Leu Leu Arg Ile Arg Pro Gly Glu Leu Gln Met Leu
                    560                     565                     570
Cys Pro Gln Glu Ala Pro Arg Ile Leu Ser Arg Leu Glu Ala Val
                    575                     580                     585
Arg Arg Met Leu Gly Ile Ser Pro
                    590
<210> 5
<211> 358
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2020010CD1

<400> 5
Met Ala Gly Pro Gly Pro Thr Phe Pro Leu His Arg Leu Val Trp
  1                 5                      10                      15
Ala Asn Arg His Arg Glu Leu Glu Ala Ala Leu His Ser His Gln
                    20                     25                      30
His Asp Ile Glu Gln Glu Asp Pro Arg Gly Arg Thr Pro Leu Glu
                    35                     40                      45
Leu Ala Val Ser Leu Gly Asn Leu Glu Ser Val Arg Val Leu Leu
                    50                     55                      60
Arg His Asn Ala Asn Val Gly Lys Glu Asn Arg Gln Gly Trp Ala
                    65                     70                      75
Val Leu Gln Glu Ala Val Ser Thr Gly Asp Pro Glu Met Val Gln
                    80                     85                      90
Leu Val Leu Gln Tyr Arg Asp Tyr Gln Arg Ala Thr Gln Arg Leu
                    95                     100                     105
Ala Gly Ile Pro Glu Leu Leu Asn Lys Leu Arg Gln Ala Pro Asp
                    110                    115                     120
Phe Tyr Val Glu Met Lys Trp Glu Phe Thr Ser Trp Val Pro Leu
                    125                    130                     135
Val Ser Lys Met Cys Pro Ser Asp Val Tyr Arg Val Trp Lys Arg
                    140                    145                     150
Gly Glu Ser Leu Arg Val Asp Thr Ser Leu Leu Gly Phe Glu His
                    155                    160                     165
```

```
Met Thr Trp Gln Arg Gly Arg Arg Ser Phe Ile Phe Lys Gly Gln
                170                 175                 180
Glu Ala Gly Ala Leu Val Met Glu Val Asp His Asp Arg Gln Val
                185                 190                 195
Val His Val Glu Thr Leu Gly Leu Thr Leu Gln Glu Pro Glu Thr
                200                 205                 210
Leu Leu Ala Ala Met Arg Pro Ser Glu Glu His Val Ala Ser Arg
                215                 220                 225
Leu Thr Ser Pro Ile Val Ser Thr His Leu Asp Thr Arg Asn Val
                230                 235                 240
Ala Phe Glu Arg Asn Lys Cys Gly Ile Trp Gly Trp Arg Ser Glu
                245                 250                 255
Lys Met Glu Thr Val Ser Gly Tyr Glu Ala Lys Val Tyr Ser Ala
                260                 265                 270
Thr Asn Val Glu Leu Val Thr Arg Thr Arg Thr Glu His Leu Ser
                275                 280                 285
Asp Gln Asp Lys Ser Arg Ser Lys Ala Gly Lys Thr Pro Phe Gln
                290                 295                 300
Ser Phe Leu Gly Met Ala Gln Gln His Ser Ser His Thr Gly Ala
                305                 310                 315
Pro Val Gln Gln Ala Ala Ser Pro Thr Asn Pro Thr Ala Ile Ser
                320                 325                 330
Pro Glu Glu Tyr Phe Asp Pro Asn Phe Ser Leu Glu Ser Arg Asn
                335                 340                 345
Ile Gly Arg Pro Ile Glu Met Ser Ser Lys Val Gln Arg
                350                 355
```

<210> 6
<211> 749
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2149037CD1

<400> 6

```
Met Ser Gly Ser His Thr Pro Ala Cys Gly Pro Phe Ser Ala Leu
  1               5                  10                  15
Thr Pro Ser Ile Trp Pro Gln Glu Ile Leu Ala Lys Tyr Thr Gln
                 20                  25                  30
Lys Glu Glu Ser Ala Glu Gln Pro Glu Phe Tyr Tyr Asp Glu Phe
                 35                  40                  45
Gly Phe Arg Val Tyr Lys Glu Glu Gly Asp Glu Pro Gly Ser Ser
                 50                  55                  60
Leu Leu Ala Asn Ser Pro Leu Met Glu Asp Ala Pro Gln Arg Leu
                 65                  70                  75
Arg Trp Gln Ala His Leu Glu Phe Thr His Asn His Asp Val Gly
                 80                  85                  90
Asp Leu Thr Trp Asp Lys Ile Ala Val Ser Leu Pro Arg Ser Glu
                 95                 100                 105
Lys Leu Arg Ser Leu Val Leu Ala Gly Ile Pro His Gly Met Arg
                110                 115                 120
Pro Gln Leu Trp Met Arg Leu Ser Gly Ala Leu Gln Lys Lys Arg
                125                 130                 135
Asn Ser Glu Leu Ser Tyr Arg Glu Ile Val Lys Asn Ser Ser Asn
                140                 145                 150
Asp Glu Thr Ile Ala Ala Lys Gln Ile Glu Lys Asp Leu Leu Arg
                155                 160                 165
Thr Met Pro Ser Asn Ala Cys Phe Ala Ser Met Gly Ser Ile Gly
                170                 175                 180
Val Pro Arg Leu Arg Arg Val Leu Arg Ala Leu Ala Trp Leu Tyr
                185                 190                 195
Pro Glu Ile Gly Tyr Cys Gln Gly Thr Gly Met Val Ala Ala Cys
                200                 205                 210
Leu Leu Leu Phe Leu Glu Glu Glu Asp Ala Phe Trp Met Met Ser
                215                 220                 225
Ala Ile Ile Glu Asp Leu Leu Pro Ala Ser Tyr Phe Ser Thr Thr
```

```
                        230                     235                     240
Leu Leu Gly Val Gln Thr Asp Gln Arg Val Leu Arg His Leu Ile
                245                     250                     255
Val Gln Tyr Leu Pro Arg Leu Asp Lys Leu Leu Gln Glu His Asp
                260                     265                     270
Ile Glu Leu Ser Leu Ile Thr Leu His Trp Phe Leu Thr Ala Phe
                275                     280                     285
Ala Ser Val Val Asp Ile Lys Leu Leu Leu Arg Ile Trp Asp Leu
                290                     295                     300
Phe Phe Tyr Glu Gly Ser Arg Val Leu Phe Gln Leu Thr Leu Gly
                305                     310                     315
Met Leu His Leu Lys Glu Glu Glu Leu Ile Gln Ser Glu Asn Ser
                320                     325                     330
Ala Ser Ile Phe Asn Thr Leu Ser Asp Ile Pro Ser Gln Met Glu
                335                     340                     345
Asp Ala Glu Leu Leu Leu Gly Val Ala Met Arg Leu Ala Gly Ser
                350                     355                     360
Leu Thr Asp Val Ala Val Glu Thr Gln Arg Arg Lys His Leu Ala
                365                     370                     375
Tyr Leu Ile Ala Asp Gln Gly Gln Leu Leu Gly Ala Gly Thr Leu
                380                     385                     390
Thr Asn Leu Ser Gln Val Val Arg Arg Arg Thr Gln Arg Arg Lys
                395                     400                     405
Ser Thr Ile Thr Ala Leu Leu Phe Gly Glu Asp Asp Leu Glu Ala
                410                     415                     420
Leu Lys Ala Lys Asn Ile Lys Gln Thr Glu Leu Val Ala Asp Leu
                425                     430                     435
Arg Glu Ala Ile Leu Arg Val Ala Arg His Phe Gln Cys Thr Asp
                440                     445                     450
Pro Lys Asn Cys Ser Val Glu Leu Thr Pro Asp Tyr Ser Met Glu
                455                     460                     465
Ser His Gln Arg Asp His Glu Asn Tyr Val Ala Cys Ser Arg Ser
                470                     475                     480
His Arg Arg Arg Ala Lys Ala Leu Leu Asp Phe Glu Arg His Asp
                485                     490                     495
Asp Asp Glu Leu Gly Phe Arg Lys Asn Asp Ile Ile Thr Ile Val
                500                     505                     510
Ser Gln Lys Asp Glu His Cys Trp Val Gly Glu Leu Asn Gly Leu
                515                     520                     525
Arg Gly Trp Phe Pro Ala Lys Phe Val Glu Val Leu Asp Glu Arg
                530                     535                     540
Ser Lys Glu Tyr Ser Ile Ala Gly Asp Asp Ser Val Thr Glu Gly
                545                     550                     555
Val Thr Asp Leu Val Arg Gly Thr Leu Cys Pro Ala Leu Lys Ala
                560                     565                     570
Leu Phe Glu His Gly Leu Lys Lys Pro Ser Leu Leu Gly Gly Ala
                575                     580                     585
Cys His Pro Trp Leu Phe Ile Glu Glu Ala Ala Gly Arg Glu Val
                590                     595                     600
Glu Arg Asp Phe Ala Ser Val Tyr Ser Arg Leu Val Leu Cys Lys
                605                     610                     615
Thr Phe Arg Leu Asp Glu Asp Gly Lys Val Leu Thr Pro Glu Glu
                620                     625                     630
Leu Leu Tyr Arg Ala Val Gln Ser Val Asn Val Thr His Asp Ala
                635                     640                     645
Val His Ala Gln Met Asp Val Lys Leu Arg Ser Leu Ile Cys Val
                650                     655                     660
Gly Leu Asn Glu Gln Val Leu His Leu Trp Leu Glu Val Leu Cys
                665                     670                     675
Ser Ser Leu Pro Thr Val Glu Lys Trp Tyr Gln Pro Trp Ser Phe
                680                     685                     690
Leu Arg Ser Pro Gly Trp Val Gln Ile Lys Cys Glu Leu Arg Val
                695                     700                     705
Leu Cys Cys Phe Ala Phe Ser Leu Ser Gln Asp Trp Glu Leu Pro
                710                     715                     720
Ala Lys Arg Glu Ala Gln Gln Pro Leu Lys Glu Gly Val Arg Asp
                725                     730                     735
```

Met Leu Val Lys His His Leu Phe Ser Trp Asp Val Asp Gly
740 745

<210> 7
<211> 139
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2162179CD1

<400> 7
Met Ala Asp Glu Lys Asp Arg Glu Glu Ile Ile Val Ala Glu Phe
1 5 10 15
His Lys Lys Ile Lys Glu Ala Phe Glu Val Phe Asp His Glu Ser
20 25 30
Asn Asn Thr Val Asp Val Arg Glu Ile Gly Thr Ile Ile Arg Ser
35 40 45
Leu Gly Cys Cys Pro Thr Glu Gly Glu Leu His Asp Leu Ile Ala
50 55 60
Glu Val Glu Glu Glu Glu Pro Thr Gly Tyr Ile Arg Phe Glu Lys
65 70 75
Phe Leu Pro Val Met Thr Glu Ile Leu Leu Glu Arg Lys Tyr Arg
80 85 90
Pro Ile Pro Glu Asp Val Leu Leu Arg Ala Phe Glu Val Leu Asp
95 100 105
Ser Ala Lys Arg Gly Phe Leu Thr Lys Asp Glu Leu Ile Lys Tyr
110 115 120
Met Thr Glu Glu Gly Lys Cys Asp Leu Leu Leu Ile Thr Met Thr
125 130 135
Tyr Val Arg Asn

<210> 8
<211> 539
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2244706CD1

<400> 8
Met Val Gly Lys Pro Val His Lys Gly Ser Glu Ser Pro Asn Ser
1 5 10 15
Phe Leu Asp Gln Glu Tyr Arg Lys Arg Phe Asn Ile Val Glu Glu
20 25 30
Asp Thr Val Leu Tyr Cys Tyr Glu Tyr Glu Lys Gly Arg Ser Ser
35 40 45
Ser Gln Gly Arg Arg Glu Ser Thr Pro Thr Tyr Gly Lys Leu Arg
50 55 60
Pro Ile Ser Met Pro Val Glu Tyr Asn Trp Val Gly Asp Tyr Glu
65 70 75
Asp Pro Asn Lys Met Lys Arg Asp Ser Arg Arg Glu Asn Ser Leu
80 85 90
Leu Arg Tyr Met Ser Asn Glu Lys Ile Ala Gln Glu Glu Tyr Met
95 100 105
Phe Gln Arg Asn Ser Lys Lys Asp Thr Gly Lys Lys Ser Lys Lys
110 115 120
Lys Gly Asp Lys Ser Asn Ser Pro Thr His Tyr Ser Leu Leu Pro
125 130 135
Ser Leu Gln Met Asp Ala Leu Arg Gln Asp Ile Met Gly Thr Pro
140 145 150
Val Pro Glu Thr Thr Leu Tyr His Thr Phe Gln Gln Ser Ser Leu
155 160 165
Gln His Lys Ser Lys Lys Lys Asn Lys Gly Pro Ile Ala Gly Lys
170 175 180
Ser Lys Arg Arg Ile Ser Cys Lys Asp Leu Gly Arg Gly Asp Cys

185 190 195
Glu Gly Trp Leu Trp Lys Lys Lys Asp Ala Lys Ser Tyr Phe Ser
200 205 210
Gln Lys Trp Lys Lys Tyr Trp Phe Val Leu Lys Asp Ala Ser Leu
215 220 225
Tyr Trp Tyr Ile Asn Glu Glu Asp Glu Lys Ala Glu Gly Phe Ile
230 235 240
Ser Leu Pro Glu Phe Lys Ile Asp Arg Ala Ser Glu Cys Arg Lys
245 250 255
Lys Tyr Ala Phe Lys Ala Cys His Pro Lys Ile Lys Ser Phe Tyr
260 265 270
Phe Ala Ala Glu His Leu Asp Asp Met Asn Arg Trp Leu Asn Arg
275 280 285
Ile Asn Met Leu Thr Ala Gly Tyr Ala Glu Arg Glu Arg Ile Lys
290 295 300
Gln Glu Gln Asp Tyr Trp Ser Glu Ser Asp Lys Glu Glu Ala Asp
305 310 315
Thr Pro Ser Thr Pro Lys Gln Asp Ser Pro Pro Pro Pro Tyr Asp
320 325 330
Thr Tyr Pro Arg Pro Pro Ser Met Ser Cys Ala Ser Pro Tyr Val
335 340 345
Glu Ala Lys His Ser Arg Leu Ser Ser Thr Glu Thr Ser Gln Ser
350 355 360
Gln Ser Ser His Glu Glu Phe Arg Gln Glu Val Thr Gly Ser Ser
365 370 375
Ala Val Ser Pro Ile Arg Lys Thr Ala Ser Gln Arg Arg Ser Trp
380 385 390
Gln Asp Leu Ile Glu Thr Pro Leu Thr Ser Ser Gly Leu His Tyr
395 400 405
Leu Gln Thr Leu Pro Leu Glu Asp Ser Val Phe Ser Asp Ser Ala
410 415 420
Ala Ile Ser Pro Glu His Arg Arg Gln Ser Thr Leu Pro Thr Gln
425 430 435
Lys Cys His Leu Gln Asp His Tyr Gly Pro Tyr Pro Leu Ala Glu
440 445 450
Ser Glu Met Met Gln Val Leu Asn Gly Asn Gly Gly Lys Pro Arg
455 460 465
Arg Phe Thr Leu Pro Arg Asp Ser Gly Phe Asn His Cys Cys Leu
470 475 480
Asn Ala Pro Val Ser Ala Cys Asp Pro Gln Asp Asp Val Gln Pro
485 490 495
Pro Glu Val Glu Glu Glu Glu Asp Asp Glu Glu Glu Ala Trp Glu
500 505 510
Ala Ala Gly Gly Asn Met Gly Glu Lys Ser Leu Phe Thr Ala Arg
515 520 525
Val Gly Arg Pro Phe Met Gln Asn Gly Ser Thr Leu Trp His
530 535

<210> 9
<211> 319
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2316805CD1

<400> 9
Met Asp Pro Gln Asn Gln His Gly Ser Gly Ser Ser Leu Val Val
1 5 10 15
Ile Gln Gln Pro Ser Leu Asp Ser Arg Gln Arg Leu Asp Tyr Glu
20 25 30
Arg Glu Ile Gln Pro Thr Ala Ile Leu Ser Leu Asp Gln Ile Lys
35 40 45
Ala Ile Arg Gly Ser Asn Glu Tyr Thr Glu Gly Pro Ser Val Val
50 55 60
Lys Arg Pro Ala Pro Arg Thr Ala Pro Arg Gln Glu Lys His Glu
65 70 75

```
Arg Thr His Glu Ile Ile Pro Ile Asn Val Asn Asn Asn Tyr Glu
                80                  85                  90
His Arg His Thr Ser His Leu Gly His Ala Val Leu Pro Ser Asn
                95                  100                 105
Ala Arg Gly Pro Ile Leu Ser Arg Ser Thr Ser Thr Gly Ser Ala
                110                 115                 120
Ala Ser Ser Gly Ser Asn Ser Ser Ala Ser Ser Glu Gln Gly Leu
                125                 130                 135
Leu Gly Arg Ser Pro Pro Thr Arg Pro Val Pro Gly His Arg Ser
                140                 145                 150
Glu Arg Ala Ile Arg Thr Gln Pro Lys Gln Leu Ile Val Asp Asp
                155                 160                 165
Leu Lys Gly Ser Leu Lys Glu Asp Leu Thr Gln His Lys Phe Ile
                170                 175                 180
Cys Glu Gln Cys Gly Lys Cys Lys Cys Gly Glu Cys Thr Ala Pro
                185                 190                 195
Arg Thr Leu Pro Ser Cys Leu Ala Cys Asn Arg Gln Cys Leu Cys
                200                 205                 210
Ser Ala Glu Ser Met Val Glu Tyr Gly Thr Cys Met Cys Leu Val
                215                 220                 225
Lys Gly Ile Phe Tyr His Cys Ser Asn Asp Asp Glu Gly Asp Ser
                230                 235                 240
Tyr Ser Asp Asn Pro Cys Ser Cys Ser Gln Ser His Cys Cys Ser
                245                 250                 255
Arg Tyr Leu Cys Met Gly Ala Met Ser Leu Phe Leu Pro Cys Leu
                260                 265                 270
Leu Cys Tyr Pro Pro Ala Lys Gly Cys Leu Lys Leu Cys Arg Arg
                275                 280                 285
Cys Tyr Asp Trp Ile His Arg Pro Gly Cys Arg Cys Lys Asn Ser
                290                 295                 300
Asn Thr Val Tyr Cys Lys Leu Glu Ser Cys Pro Ser Arg Gly Gln
                305                 310                 315
Gly Lys Pro Ser

<210> 10
<211> 747
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2320010CD1

<400> 10
Met Gly Lys Arg Asn Ile Ala Arg Val His Asp Ala Trp Leu Ser
  1             5                   10                  15
Lys His Phe Gly Ile Asp Arg Lys Ser Gln Thr Met Pro Ala Leu
                20                  25                  30
Arg Asn Arg Ser Gly Val Met Gln Ala Arg Leu Gln His Leu Ser
                35                  40                  45
Ser Leu Glu Ser Ser Phe Thr Leu Asn His Ser Ser Thr Thr Thr
                50                  55                  60
Glu Ala Asp Ile Phe His Gln Ala Leu Leu Ala Ala Asn Thr Ala
                65                  70                  75
Thr Glu Val Ser Leu Thr Val Leu Asp Thr Ile Ser Phe Phe Thr
                80                  85                  90
Gln Cys Phe Lys Thr Gln Leu Leu Asn Asn Asp Gly His Asn Pro
                95                  100                 105
Leu Met Lys Lys Val Phe Asp Ile His Leu Ala Phe Leu Lys Asn
                110                 115                 120
Gly Gln Ser Glu Val Ser Leu Lys His Val Phe Ala Ser Leu Arg
                125                 130                 135
Ala Phe Ile Ser Lys Phe Pro Ser Ala Phe Phe Lys Gly Arg Val
                140                 145                 150
Asn Met Cys Ala Ala Phe Cys Tyr Glu Val Leu Lys Cys Cys Thr
                155                 160                 165
Ser Lys Ile Ser Ser Thr Arg Asn Glu Ala Ser Ala Leu Leu Tyr
```

170 175 180
Leu Leu Met Arg Asn Asn Phe Glu Tyr Thr Lys Arg Lys Thr Phe
185 190 195
Leu Arg Thr His Leu Gln Ile Ile Ile Ala Val Ser Gln Leu Ile
200 205 210
Ala Asp Val Ala Leu Ser Gly Gly Ser Arg Phe Gln Glu Ser Leu
215 220 225
Phe Ile Ile Asn Asn Phe Ala Asn Ser Asp Arg Pro Met Lys Ala
230 235 240
Thr Ala Phe Pro Ala Glu Val Lys Asp Leu Thr Lys Arg Ile Arg
245 250 255
Thr Val Leu Met Ala Thr Ala Gln Met Lys Glu His Glu Lys Asp
260 265 270
Pro Glu Met Leu Ile Asp Leu Gln Tyr Ser Leu Ala Lys Ser Tyr
275 280 285
Ala Ser Thr Pro Glu Leu Arg Lys Thr Trp Leu Asp Ser Met Ala
290 295 300
Lys Ile His Val Lys Asn Gly Asp Phe Ser Glu Ala Ala Met Cys
305 310 315
Tyr Val His Val Ala Ala Leu Val Ala Glu Phe Leu His Arg Lys
320 325 330
Lys Leu Phe Pro Asn Gly Cys Ser Ala Phe Lys Lys Ile Thr Pro
335 340 345
Asn Ile Asp Glu Glu Gly Ala Met Lys Glu Asp Ala Gly Met Met
350 355 360
Asp Val His Tyr Ser Glu Glu Val Leu Leu Glu Leu Leu Glu Gln
365 370 375
Cys Val Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Ile Ile Ser
380 385 390
Glu Ile Ser Lys Leu Ile Val Pro Ile Tyr Glu Lys Arg Arg Glu
395 400 405
Phe Glu Lys Leu Thr Gln Val Tyr Arg Thr Leu His Gly Ala Tyr
410 415 420
Thr Lys Ile Leu Glu Val Met His Thr Lys Lys Arg Leu Leu Gly
425 430 435
Thr Phe Phe Arg Val Ala Phe Tyr Gly Gln Ser Phe Phe Glu Glu
440 445 450
Glu Asp Gly Lys Glu Tyr Ile Tyr Lys Glu Pro Lys Leu Thr Gly
455 460 465
Leu Ser Glu Ile Ser Leu Arg Leu Val Lys Leu Tyr Gly Glu Lys
470 475 480
Phe Gly Thr Glu Asn Val Lys Ile Ile Gln Asp Ser Asp Lys Val
485 490 495
Asn Ala Lys Glu Leu Asp Pro Lys Tyr Ala His Ile Gln Val Thr
500 505 510
Tyr Val Lys Pro Tyr Phe Asp Asp Lys Glu Leu Thr Glu Arg Lys
515 520 525
Thr Glu Phe Glu Arg Asn His Asn Ile Ser Arg Phe Val Phe Glu
530 535 540
Ala Pro Tyr Thr Leu Ser Gly Lys Lys Gln Gly Cys Ile Glu Glu
545 550 555
Gln Cys Lys Arg Arg Thr Ile Leu Thr Thr Ser Asn Ser Phe Pro
560 565 570
Tyr Val Lys Lys Arg Ile Pro Ile Asn Cys Glu Gln Gln Ile Asn
575 580 585
Leu Lys Pro Ile Asp Val Ala Thr Asp Glu Ile Lys Asp Lys Thr
590 595 600
Ala Glu Leu Gln Lys Leu Cys Ser Ser Thr Asp Val Asp Met Ile
605 610 615
Gln Leu Gln Leu Lys Leu Gln Gly Cys Val Ser Val Gln Val Asn
620 625 630
Ala Gly Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asn Asp Ser Gln
635 640 645
Ala Ser Lys Tyr Pro Pro Lys Lys Val Ser Glu Leu Lys Asp Met
650 655 660
Phe Arg Lys Phe Ile Gln Ala Cys Ser Ile Ala Leu Glu Leu Asn
665 670 675

```
Glu Arg Leu Ile Lys Glu Asp Gln Val Glu Tyr His Glu Gly Leu
                680                 685                 690
Lys Ser Asn Phe Arg Asp Met Val Lys Glu Leu Ser Asp Ile Ile
                695                 700                 705
His Glu Gln Ile Leu Gln Glu Asp Thr Met His Ser Pro Trp Met
                710                 715                 720
Ser Asn Thr Leu His Val Phe Cys Ala Ile Ser Gly Thr Ser Ser
                725                 730                 735
Asp Arg Gly Tyr Gly Ser Pro Arg Tyr Ala Glu Val
                740                 745
```

<210> 11
<211> 266
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2564901CD1

<400> 11

```
Met Gln Gly Ser Thr Arg Arg Met Gly Val Met Thr Asp Val His
  1             5                   10                  15
Arg Arg Phe Leu Gln Leu Leu Met Thr His Gly Val Leu Glu Glu
                20                  25                  30
Trp Asp Val Lys Arg Leu Gln Thr His Cys Tyr Lys Val His Asp
                35                  40                  45
Arg Asn Ala Thr Val Asp Lys Leu Glu Asp Phe Ile Asn Asn Ile
                50                  55                  60
Asn Ser Val Leu Glu Ser Leu Tyr Ile Glu Ile Lys Arg Gly Val
                65                  70                  75
Thr Glu Asp Asp Gly Arg Pro Ile Tyr Ala Leu Val Asn Leu Ala
                80                  85                  90
Thr Thr Ser Ile Ser Lys Met Ala Thr Asp Phe Ala Glu Asn Glu
                95                  100                 105
Leu Asp Leu Phe Arg Lys Ala Leu Glu Leu Ile Ile Asp Ser Glu
                110                 115                 120
Thr Gly Phe Ala Ser Ser Thr Asn Ile Leu Asn Leu Val Asp Gln
                125                 130                 135
Leu Lys Gly Lys Lys Met Arg Lys Lys Glu Ala Glu Gln Val Leu
                140                 145                 150
Gln Lys Phe Val Gln Asn Lys Trp Leu Ile Glu Lys Glu Gly Glu
                155                 160                 165
Phe Thr Leu His Gly Arg Ala Ile Leu Glu Met Glu Gln Tyr Ile
                170                 175                 180
Arg Glu Thr Tyr Pro Asp Ala Val Lys Ile Cys Asn Ile Cys His
                185                 190                 195
Ser Leu Leu Ile Gln Gly Gln Ser Cys Glu Thr Cys Gly Ile Arg
                200                 205                 210
Met His Leu Pro Cys Val Ala Lys Tyr Phe Gln Ser Asn Ala Glu
                215                 220                 225
Pro Arg Cys Pro His Cys Asn Asp Tyr Trp Pro His Glu Ile Pro
                230                 235                 240
Lys Val Phe Asp Pro Glu Lys Glu Arg Glu Ser Gly Val Leu Lys
                245                 250                 255
Ser Asn Lys Lys Ser Leu Arg Ser Arg Gln His
                260                 265
```

<210> 12
<211> 345
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2615168CD1

<400> 12

```
Met Ser Val Thr Gly Gly Lys Met Ala Pro Ser Leu Thr Gln Glu
```

```
  1               5                   10                  15
Ile Leu Ser His Leu Gly Leu Ala Ser Lys Thr Ala Ala Trp Gly
                 20                  25                  30
Thr Leu Gly Thr Leu Arg Thr Phe Leu Asn Phe Ser Val Asp Lys
                 35                  40                  45
Asp Ala Gln Arg Leu Leu Arg Ala Ile Thr Gly Gln Gly Val Asp
                 50                  55                  60
Arg Ser Ala Ile Val Asp Val Leu Thr Asn Arg Ser Arg Glu Gln
                 65                  70                  75
Arg Gln Leu Ile Ser Arg Asn Phe Gln Glu Arg Thr Gln Gln Asp
                 80                  85                  90
Leu Met Lys Ser Leu Gln Ala Ala Leu Ser Gly Asn Leu Glu Arg
                 95                 100                 105
Ile Val Met Ala Leu Leu Gln Pro Thr Ala Gln Phe Asp Ala Gln
                110                 115                 120
Glu Leu Arg Thr Ala Leu Lys Ala Ser Asp Ser Ala Val Asp Val
                125                 130                 135
Ala Ile Glu Ile Leu Ala Thr Arg Thr Pro Pro Gln Leu Gln Glu
                140                 145                 150
Cys Leu Ala Val Tyr Lys His Asn Phe Gln Val Glu Ala Val Asp
                155                 160                 165
Asp Ile Thr Ser Glu Thr Ser Gly Ile Leu Gln Asp Leu Leu Leu
                170                 175                 180
Ala Leu Ala Lys Gly Gly Arg Asp Ser Tyr Ser Gly Ile Ile Asp
                185                 190                 195
Tyr Asn Leu Ala Glu Gln Asp Val Gln Ala Leu Gln Arg Ala Glu
                200                 205                 210
Gly Pro Ser Arg Glu Glu Thr Trp Val Pro Val Phe Thr Gln Arg
                215                 220                 225
Asn Pro Glu His Leu Ile Arg Val Phe Asp Gln Tyr Gln Arg Ser
                230                 235                 240
Thr Gly Gln Glu Leu Glu Glu Ala Val Gln Asn Arg Phe His Gly
                245                 250                 255
Asp Ala Gln Val Ala Leu Leu Gly Leu Ala Ser Val Ile Lys Asn
                260                 265                 270
Thr Pro Leu Tyr Phe Ala Asp Lys Leu His Gln Ala Leu Gln Glu
                275                 280                 285
Thr Glu Pro Asn Tyr Gln Val Leu Ile Arg Ile Leu Ile Ser Arg
                290                 295                 300
Cys Glu Thr Asp Leu Leu Ser Ile Arg Ala Glu Phe Arg Lys Lys
                305                 310                 315
Phe Gly Lys Ser Leu Tyr Ser Ser Leu Gln Asp Ala Val Lys Gly
                320                 325                 330
Asp Cys Gln Ser Ala Leu Leu Ala Leu Cys Arg Ala Glu Asp Met
                335                 340                 345

<210> 13
<211> 437
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2658329CD1

<400> 13
Met Glu Lys Glu Leu Arg Ser Thr Ile Leu Phe Asn Ala Tyr Lys
  1               5                  10                  15
Lys Glu Ile Phe Thr Thr Asn Asn Gly Tyr Lys Ser Met Gln Lys
                 20                  25                  30
Lys Leu Arg Ser Asn Trp Lys Ile Gln Ser Leu Lys Asp Glu Ile
                 35                  40                  45
Thr Ser Glu Lys Leu Asn Gly Val Lys Leu Trp Ile Thr Ala Gly
                 50                  55                  60
Pro Arg Glu Lys Phe Thr Ala Ala Glu Phe Glu Ile Leu Lys Lys
                 65                  70                  75
Tyr Leu Asp Thr Gly Gly Asp Val Phe Val Met Leu Gly Glu Gly
```

```
                    80                                      85                                      90
Gly Glu Ser Arg Phe Asp Thr Asn Ile Asn Phe Leu Leu Glu Glu
                95                  100                 105
Tyr Gly Ile Met Val Asn Asn Asp Ala Val Val Arg Asn Val Tyr
                110                 115                 120
His Lys Tyr Phe His Pro Lys Glu Ala Leu Val Ser Ser Gly Val
                125                 130                 135
Leu Asn Arg Glu Ile Ser Arg Ala Ala Gly Lys Ala Val Pro Gly
                140                 145                 150
Ile Ile Asp Glu Glu Ser Ser Gly Asn Asn Ala Gln Ala Leu Thr
                155                 160                 165
Phe Val Tyr Pro Phe Gly Ala Thr Leu Ser Val Met Lys Pro Ala
                170                 175                 180
Val Ala Val Leu Ser Thr Gly Ser Val Cys Phe Pro Leu Asn Arg
                185                 190                 195
Pro Ile Leu Ala Phe Tyr His Ser Lys Asn Gln Gly Gly Lys Leu
                200                 205                 210
Ala Val Leu Gly Ser Cys His Met Phe Ser Asp Gln Tyr Leu Asp
                215                 220                 225
Lys Glu Glu Asn Ser Lys Ile Met Asp Val Val Phe Gln Trp Leu
                230                 235                 240
Thr Thr Gly Asp Ile His Leu Asn Gln Ile Asp Ala Glu Asp Pro
                245                 250                 255
Glu Ile Ser Asp Tyr Met Met Leu Pro Tyr Thr Ala Thr Leu Ser
                260                 265                 270
Lys Arg Asn Arg Glu Cys Leu Gln Glu Ser Asp Glu Ile Pro Arg
                275                 280                 285
Asp Phe Thr Thr Leu Phe Asp Leu Ser Ile Phe Gln Leu Asp Thr
                290                 295                 300
Thr Ser Phe His Ser Val Ile Glu Ala His Glu Gln Leu Asn Val
                305                 310                 315
Lys His Glu Pro Leu Gln Leu Ile Gln Pro Gln Phe Glu Thr Pro
                320                 325                 330
Leu Pro Thr Leu Gln Pro Ala Val Phe Pro Pro Ser Phe Arg Glu
                335                 340                 345
Leu Pro Pro Pro Pro Leu Glu Leu Phe Asp Leu Asp Glu Thr Phe
                350                 355                 360
Ser Ser Glu Lys Ala Arg Leu Ala Gln Ile Thr Asn Lys Cys Thr
                365                 370                 375
Glu Glu Asp Leu Glu Phe Tyr Val Arg Lys Cys Gly Asp Ile Leu
                380                 385                 390
Gly Val Thr Ser Lys Leu Pro Lys Asp Gln Gln Asp Ala Lys His
                395                 400                 405
Ile Leu Glu His Val Phe Phe Gln Val Val Glu Phe Lys Lys Leu
                410                 415                 420
Asn Gln Glu His Asp Ile Asp Thr Ser Glu Thr Ala Phe Gln Asn
                425                 430                 435
Asn Phe

<210> 14
<211> 441
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2708944CD1

<400> 14
Met Val His Ile Lys Lys Gly Glu Leu Thr Gln Glu Glu Lys Glu
 1              5                   10                  15
Leu Leu Glu Val Ile Gly Lys Gly Thr Val Gln Glu Ala Gly Thr
                20                  25                  30
Leu Leu Ser Ser Lys Asn Val Arg Val Asn Cys Leu Asp Glu Asn
                35                  40                  45
Gly Met Thr Pro Leu Met His Ala Ala Tyr Lys Gly Lys Leu Asp
                50                  55                  60
```

Met Cys Lys Leu Leu Leu Arg His Gly Ala Asp Val Asn Cys His
65 70 75
Gln His Glu His Gly Tyr Thr Ala Leu Met Phe Ala Ala Leu Ser
80 85 90
Gly Asn Lys Asp Ile Thr Trp Val Met Leu Glu Ala Gly Ala Glu
95 100 105
Thr Asp Val Val Asn Ser Val Gly Arg Thr Ala Ala Gln Met Ala
110 115 120
Ala Phe Val Gly Gln His Asp Cys Val Thr Ile Ile Asn Asn Phe
125 130 135
Phe Pro Arg Glu Arg Leu Asp Tyr Tyr Thr Lys Pro Gln Gly Leu
140 145 150
Asp Lys Glu Pro Lys Leu Pro Pro Lys Leu Ala Gly Pro Leu His
155 160 165
Lys Ile Ile Thr Thr Thr Asn Leu His Pro Val Lys Ile Val Met
170 175 180
Leu Val Asn Glu Asn Pro Leu Leu Thr Glu Glu Ala Ala Leu Asn
185 190 195
Lys Cys Tyr Arg Val Met Asp Leu Ile Cys Glu Lys Cys Met Lys
200 205 210
Gln Arg Asp Met Asn Glu Val Leu Ala Met Lys Met His Tyr Ile
215 220 225
Ser Cys Ile Phe Gln Lys Cys Ile Asn Phe Leu Lys Asp Gly Glu
230 235 240
Asn Lys Leu Asp Thr Leu Ile Lys Ser Leu Leu Lys Gly Arg Ala
245 250 255
Ser Asp Gly Phe Pro Val Tyr Gln Glu Lys Ile Ile Arg Glu Ser
260 265 270
Ile Arg Lys Phe Pro Tyr Cys Glu Ala Thr Leu Leu Gln Gln Leu
275 280 285
Val Arg Ser Ile Ala Pro Val Glu Ile Gly Ser Asp Pro Thr Ala
290 295 300
Phe Ser Val Leu Thr Gln Ala Ile Thr Gly Gln Val Gly Phe Val
305 310 315
Asp Val Glu Phe Cys Thr Thr Cys Gly Glu Lys Gly Ala Ser Lys
320 325 330
Arg Cys Ser Val Cys Lys Met Val Ile Tyr Cys Asp Gln Thr Cys
335 340 345
Gln Lys Thr His Trp Phe Thr His Lys Lys Ile Cys Lys Asn Leu
350 355 360
Lys Asp Ile Tyr Glu Lys Gln Gln Leu Glu Ala Ala Lys Glu Lys
365 370 375
Arg Gln Glu Glu Asn His Gly Lys Leu Asp Val Asn Ser Asn Cys
380 385 390
Val Asn Glu Glu Gln Pro Glu Ala Glu Val Gly Ile Ser Gln Lys
395 400 405
Asp Ser Asn Pro Glu Asp Ser Gly Glu Gly Lys Lys Glu Ser Leu
410 415 420
Glu Ser Glu Ala Glu Leu Glu Gly Leu Gln Asp Ala Pro Ala Gly
425 430 435
Pro Gln Val Ser Glu Glu
440

<210> 15
<211> 487
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3315012CD1

<400> 15
Met Leu Arg Ala Pro Gly Cys Leu Leu Arg Thr Ser Val Ala Pro
1 5 10 15
Ala Ala Ala Leu Ala Ala Ala Leu Leu Ser Ser Leu Ala Arg Cys
20 25 30
Ser Leu Leu Glu Pro Arg Asp Pro Val Ala Ser Ser Leu Ser Pro

```
                    35                       40                       45
Tyr Phe Gly Thr Lys Thr Arg Tyr Glu Asp Val Asn Pro Val Leu
                    50                       55                       60
Leu Ser Gly Pro Glu Ala Pro Trp Arg Asp Pro Glu Leu Leu Glu
                    65                       70                       75
Gly Thr Cys Thr Pro Val Gln Leu Val Ala Leu Ile Arg His Gly
                    80                       85                       90
Thr Arg Tyr Pro Thr Val Lys Gln Ile Arg Lys Leu Arg Gln Leu
                    95                      100                      105
His Gly Leu Leu Gln Ala Arg Gly Ser Arg Asp Gly Gly Ala Ser
                   110                      115                      120
Ser Thr Gly Ser Arg Asp Leu Gly Ala Ala Leu Ala Asp Trp Pro
                   125                      130                      135
Leu Trp Tyr Ala Asp Trp Met Asp Gly Gln Leu Val Glu Lys Gly
                   140                      145                      150
Arg Gln Asp Met Arg Gln Leu Ala Leu Arg Leu Ala Ser Leu Phe
                   155                      160                      165
Pro Val Leu Phe Ser Arg Glu Asn Tyr Gly Arg Leu Arg Leu Ile
                   170                      175                      180
Thr Ser Ser Lys His Arg Cys Met Asp Ser Ser Ala Ala Phe Leu
                   185                      190                      195
Gln Gly Leu Trp Gln His Tyr His Pro Gly Leu Pro Pro Pro Asp
                   200                      205                      210
Val Ala Asp Met Glu Phe Gly Pro Pro Thr Val Asn Asp Lys Leu
                   215                      220                      225
Met Arg Phe Phe Asp His Cys Glu Lys Phe Leu Thr Glu Val Glu
                   230                      235                      240
Lys Asn Ala Thr Ala Leu Tyr His Val Glu Ala Phe Lys Thr Gly
                   245                      250                      255
Pro Glu Met Gln Asn Ile Leu Lys Lys Val Ala Ala Thr Leu Gln
                   260                      265                      270
Val Pro Val Asn Asp Leu Asn Ala Asp Leu Ile Gln Val Ala Phe
                   275                      280                      285
Phe Thr Cys Ser Phe Asp Leu Ala Ile Lys Gly Val Lys Ser Pro
                   290                      295                      300
Trp Cys Asp Val Phe Asp Ile Asp Asp Ala Lys Val Leu Glu Tyr
                   305                      310                      315
Leu Asn Asp Leu Lys Gln Tyr Trp Lys Arg Gly Tyr Gly Tyr Thr
                   320                      325                      330
Ile Asn Ser Arg Ser Ser Cys Thr Leu Phe Gln Asp Ile Phe Gln
                   335                      340                      345
His Leu Asp Lys Ala Val Glu Gln Lys Gln Arg Ser Gln Pro Ile
                   350                      355                      360
Ser Ser Pro Val Ile Leu Gln Phe Gly His Ala Glu Thr Leu Leu
                   365                      370                      375
Pro Leu Leu Ser Leu Met Gly Tyr Phe Lys Asp Lys Glu Pro Leu
                   380                      385                      390
Thr Ala Tyr Asn Tyr Lys Lys Gln Met His Arg Lys Phe Arg Ser
                   395                      400                      405
Gly Leu Ile Val Pro Tyr Ala Ser Asn Leu Ile Phe Val Leu Tyr
                   410                      415                      420
His Cys Glu Asn Ala Lys Thr Pro Lys Glu Gln Phe Arg Val Gln
                   425                      430                      435
Met Leu Leu Asn Glu Lys Val Leu Pro Leu Ala Tyr Ser Gln Glu
                   440                      445                      450
Thr Val Ser Phe Tyr Glu Asp Leu Lys Asn His Tyr Lys Asp Ile
                   455                      460                      465
Leu Gln Ser Cys Gln Thr Ser Glu Glu Cys Glu Leu Ala Arg Ala
                   470                      475                      480
Asn Ser Thr Ser Asp Glu Leu
                   485
```

<210> 16
<211> 282
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 4155412CD1

<400> 16
```
Met Val Leu Gly Lys Val Lys Ser Leu Thr Ile Ser Phe Asp Cys
  1               5                  10                  15
Leu Asn Asp Ser Asn Val Pro Val Tyr Ser Ser Gly Asp Thr Val
                 20                  25                  30
Ser Gly Arg Val Asn Leu Glu Val Thr Gly Glu Ile Arg Val Lys
                 35                  40                  45
Ser Leu Lys Ile His Ala Arg Gly His Ala Lys Val Arg Trp Thr
                 50                  55                  60
Glu Ser Arg Asn Ala Gly Ser Asn Thr Ala Tyr Thr Gln Asn Tyr
                 65                  70                  75
Thr Glu Glu Val Glu Tyr Phe Asn His Lys Asp Ile Leu Ile Gly
                 80                  85                  90
His Glu Arg Asp Asp Asp Asn Ser Glu Glu Gly Phe His Thr Ile
                 95                 100                 105
His Ser Gly Arg His Glu Tyr Ala Phe Ser Phe Glu Leu Pro Gln
                110                 115                 120
Thr Pro Leu Ala Thr Ser Phe Glu Gly Arg His Gly Ser Val Arg
                125                 130                 135
Tyr Trp Val Lys Ala Glu Leu His Arg Pro Trp Leu Leu Pro Val
                140                 145                 150
Lys Leu Lys Lys Glu Phe Thr Val Phe Glu His Ile Asp Ile Asn
                155                 160                 165
Thr Pro Ser Leu Leu Ser Pro Gln Ala Gly Thr Lys Glu Lys Thr
                170                 175                 180
Leu Cys Cys Trp Phe Cys Thr Ser Gly Pro Ile Ser Leu Ser Ala
                185                 190                 195
Lys Ile Glu Arg Lys Gly Tyr Thr Pro Gly Glu Ser Ile Gln Ile
                200                 205                 210
Phe Ala Glu Ile Glu Asn Cys Ser Ser Arg Met Val Val Pro Arg
                215                 220                 225
Gln Pro Phe Thr Lys His Arg Pro Ser Ile Ala Lys Gly Lys Leu
                230                 235                 240
Arg Glu Leu Asn Ser Leu Trp Leu Thr Cys Val Gly Asn Ser Leu
                245                 250                 255
Thr Ser Gly Lys Asn Arg Asp Val Glu Met Ala Ser Leu Leu Lys
                260                 265                 270
Ile Ser Asn Ser Phe Pro Pro Ser Asn Ala Ser Asn
                275                 280
```
<210> 17
<211> 581
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4831840CD1

<400> 17
```
Met Ala Val Ala Gly Ala Val Ser Gly Glu Pro Leu Val His Trp
  1               5                  10                  15
Cys Thr Gln Gln Leu Arg Lys Thr Phe Gly Leu Asp Val Ser Glu
                 20                  25                  30
Glu Ile Ile Gln Tyr Val Leu Ser Ile Glu Ser Ala Glu Glu Ile
                 35                  40                  45
Arg Glu Tyr Val Thr Asp Leu Leu Gln Gly Asn Glu Gly Lys Lys
                 50                  55                  60
Gly Gln Phe Ile Glu Glu Leu Ile Thr Lys Trp Gln Lys Asn Asp
                 65                  70                  75
Gln Glu Leu Ile Ser Asp Pro Leu Gln Gln Cys Phe Lys Lys Asp
                 80                  85                  90
Glu Ile Leu Asp Gly Gln Lys Ser Gly Asp His Leu Lys Arg Gly
                 95                 100                 105
Arg Lys Lys Gly Arg Asn Arg Gln Glu Val Pro Ala Phe Thr Glu
```

```
                110                     115                     120
Pro Asp Thr Thr Ala Glu Val Lys Thr Pro Phe Asp Leu Ala Lys
                125                     130                     135
Ala Gln Glu Asn Ser Asn Ser Val Lys Lys Lys Thr Lys Phe Val
                140                     145                     150
Asn Leu Tyr Thr Arg Glu Gly Gln Asp Arg Leu Ala Val Leu Leu
                155                     160                     165
Pro Gly Arg His Pro Cys Asp Cys Leu Gly Gln Lys His Lys Leu
                170                     175                     180
Ile Asn Asn Cys Leu Ile Cys Gly Arg Ile Val Cys Glu Gln Glu
                185                     190                     195
Gly Ser Gly Pro Cys Leu Phe Cys Gly Thr Leu Val Cys Thr His
                200                     205                     210
Glu Glu Gln Asp Ile Leu Gln Arg Asp Ser Asn Lys Ser Gln Lys
                215                     220                     225
Leu Leu Lys Lys Leu Met Ser Gly Val Glu Asn Ser Gly Lys Val
                230                     235                     240
Asp Ile Ser Thr Lys Asp Leu Leu Pro His Gln Glu Leu Arg Ile
                245                     250                     255
Lys Ser Gly Leu Glu Lys Ala Ile Lys His Lys Asp Lys Leu Leu
                260                     265                     270
Glu Phe Asp Arg Thr Ser Ile Arg Arg Thr Gln Val Ile Asp Asp
                275                     280                     285
Glu Ser Asp Tyr Phe Ala Ser Asp Ser Asn Gln Trp Leu Ser Lys
                290                     295                     300
Leu Glu Arg Glu Thr Leu Gln Lys Arg Glu Glu Glu Leu Arg Glu
                305                     310                     315
Leu Arg His Ala Ser Arg Leu Ser Lys Lys Val Thr Ile Asp Phe
                320                     325                     330
Ala Gly Arg Lys Ile Leu Glu Glu Glu Asn Ser Leu Ala Glu Tyr
                335                     340                     345
His Ser Arg Leu Asp Glu Thr Ile Gln Ala Ile Ala Asn Gly Thr
                350                     355                     360
Leu Asn Gln Pro Leu Thr Lys Leu Asp Arg Ser Ser Glu Glu Pro
                365                     370                     375
Leu Gly Val Leu Val Asn Pro Asn Met Tyr Gln Ser Pro Pro Gln
                380                     385                     390
Trp Val Asp His Thr Gly Ala Ala Ser Gln Lys Lys Ala Phe Arg
                395                     400                     405
Ser Ser Gly Phe Gly Leu Glu Phe Asn Ser Phe Gln His Gln Leu
                410                     415                     420
Arg Ile Gln Asp Gln Glu Phe Gln Glu Gly Phe Asp Gly Gly Trp
                425                     430                     435
Cys Leu Ser Val His Gln Pro Trp Ala Ser Leu Leu Val Arg Gly
                440                     445                     450
Ile Lys Arg Val Glu Gly Arg Ser Trp Tyr Thr Pro His Arg Gly
                455                     460                     465
Arg Leu Trp Ile Ala Ala Thr Ala Lys Lys Pro Ser Pro Gln Glu
                470                     475                     480
Val Ser Glu Leu Gln Ala Thr Tyr Arg Leu Leu Arg Gly Lys Asp
                485                     490                     495
Val Glu Phe Pro Asn Asp Tyr Pro Ser Gly Cys Leu Leu Gly Cys
                500                     505                     510
Val Asp Leu Ile Asp Cys Leu Ser Gln Lys Gln Phe Lys Glu Gln
                515                     520                     525
Phe Pro Asp Ile Ser Gln Glu Ser Asp Ser Pro Phe Val Phe Ile
                530                     535                     540
Cys Lys Asn Pro Gln Glu Met Val Val Lys Phe Pro Ile Lys Gly
                545                     550                     555
Asn Pro Lys Ile Trp Lys Leu Asp Ser Lys Ile His Gln Gly Ala
                560                     565                     570
Lys Lys Gly Leu Met Lys Gln Asn Lys Ala Val
                575                     580
```

<210> 18
<211> 530
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5676581CD1

<400> 18
Met Thr Thr Arg Pro Thr Ala Val Lys Ala Thr Gly Gly Leu Cys
1 5 10 15
Leu Leu Gly Ala Tyr Ala Asp Ser Asp Asp Asp Asp Asn Asp Val
20 25 30
Ser Glu Lys Leu Ala Gln Ser Lys Glu Thr Asn Gly Asn Gln Ser
35 40 45
Thr Asp Ile Asp Ser Thr Leu Ala Asn Phe Leu Ala Glu Ile Asp
50 55 60
Ala Ile Thr Ala Pro Gln Pro Ala Ala Pro Val Gly Ala Ser Ala
65 70 75
Pro Pro Pro Thr Pro Pro Arg Pro Glu Pro Lys Glu Ala Ala Thr
80 85 90
Ser Thr Leu Ser Ser Ser Thr Ser Asn Gly Thr Asp Ser Thr Gln
95 100 105
Thr Ser Gly Trp Gln Tyr Asp Thr Gln Cys Ser Leu Ala Gly Val
110 115 120
Gly Ile Glu Met Gly Asp Trp Gln Glu Val Trp Asp Glu Asn Thr
125 130 135
Gly Cys Tyr Tyr Tyr Trp Asn Thr Gln Thr Asn Glu Val Thr Trp
140 145 150
Glu Leu Pro Gln Tyr Leu Ala Thr Gln Val Gln Gly Leu Gln His
155 160 165
Tyr Gln Pro Ser Ser Val Pro Gly Ala Glu Thr Ser Phe Val Val
170 175 180
Asn Thr Asp Ile Tyr Ser Lys Glu Lys Thr Ile Ser Val Ser Ser
185 190 195
Ser Lys Ser Gly Pro Val Ile Ala Lys Arg Glu Val Lys Lys Glu
200 205 210
Val Asn Glu Gly Ile Gln Ala Leu Ser Asn Ser Glu Glu Glu Lys
215 220 225
Lys Gly Val Ala Ala Ser Leu Leu Ala Pro Leu Leu Pro Glu Gly
230 235 240
Ile Lys Glu Glu Glu Glu Arg Trp Arg Arg Lys Val Ile Cys Lys
245 250 255
Glu Glu Pro Val Ser Glu Val Lys Glu Thr Ser Thr Thr Val Glu
260 265 270
Glu Ala Thr Thr Ile Val Lys Pro Gln Glu Ile Met Leu Asp Asn
275 280 285
Ile Glu Asp Pro Ser Gln Glu Asp Leu Cys Ser Val Val Gln Ser
290 295 300
Gly Glu Ser Glu Glu Glu Glu Glu Gln Asp Thr Leu Glu Leu Glu
305 310 315
Leu Val Leu Glu Arg Lys Lys Ala Glu Leu Arg Ala Leu Glu Glu
320 325 330
Gly Asp Gly Ser Val Ser Gly Ser Ser Pro Arg Ser Asp Ile Ser
335 340 345
Gln Pro Ala Ser Gln Asp Gly Met Arg Arg Leu Met Ser Lys Arg
350 355 360
Gly Lys Trp Lys Met Phe Val Arg Ala Thr Ser Pro Glu Ser Thr
365 370 375
Ser Arg Ser Ser Ser Lys Thr Gly Arg Asp Thr Pro Glu Asn Gly
380 385 390
Glu Thr Ala Ile Gly Ala Glu Asn Ser Glu Lys Ile Asp Glu Asn
395 400 405
Ser Asp Lys Glu Met Glu Val Glu Glu Ser Pro Glu Lys Ile Lys
410 415 420
Val Gln Thr Thr Pro Lys Val Glu Glu Glu Gln Asp Leu Lys Phe
425 430 435
Gln Ile Gly Glu Leu Ala Asn Thr Leu Thr Ser Lys Phe Glu Phe
440 445 450
Leu Gly Ile Asn Arg Gln Ser Ile Ser Asn Phe His Val Leu Leu
455 460 465

```
Leu Gln Thr Glu Thr Arg Ile Ala Asp Trp Arg Glu Gly Ala Leu
                470                 475                 480
Asn Gly Asn Tyr Leu Lys Arg Lys Leu Gln Asp Ala Ala Glu Gln
                485                 490                 495
Leu Lys Gln Tyr Glu Ile Asn Ala Thr Pro Lys Gly Trp Ser Cys
                500                 505                 510
His Trp Asp Arg Tyr Ala Leu Phe Ser Pro Phe His Leu Ser Pro
                515                 520                 525
Leu Thr Ser Gln Thr
                530
<210> 19
<211> 475
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 034159CD1

<400> 19
Met Gln Lys Ser Thr Asn Ser Asp Thr Ser Val Glu Thr Leu Asn
  1               5                  10                  15
Ser Thr Arg Gln Gly Thr Gly Ala Val Gln Met Arg Ile Lys Asn
                 20                  25                  30
Ala Asn Ser His His Asp Arg Leu Ser Gln Ser Lys Ser Met Ile
                 35                  40                  45
Leu Thr Asp Val Gly Lys Val Thr Glu Pro Ile Ser Arg His Arg
                 50                  55                  60
Arg Asn His Ser Gln His Ile Leu Lys Asp Val Ile Pro Pro Leu
                 65                  70                  75
Glu Gln Leu Met Val Glu Lys Glu Gly Tyr Leu Gln Lys Ala Lys
                 80                  85                  90
Ile Ala Asp Gly Gly Lys Lys Leu Arg Lys Asn Trp Ser Thr Ser
                 95                 100                 105
Trp Ile Val Leu Ser Ser Arg Arg Ile Glu Phe Tyr Lys Glu Ser
                110                 115                 120
Lys Gln Gln Ala Leu Ser Asn Met Lys Thr Gly His Lys Pro Glu
                125                 130                 135
Ser Val Asp Leu Cys Gly Ala His Ile Glu Trp Ala Lys Glu Lys
                140                 145                 150
Ser Ser Arg Lys Asn Val Phe Gln Ile Thr Thr Val Ser Gly Asn
                155                 160                 165
Glu Phe Leu Leu Gln Ser Asp Ile Asp Phe Ile Ile Leu Asp Trp
                170                 175                 180
Phe His Ala Ile Lys Asn Ala Ile Asp Arg Leu Pro Lys Asp Ser
                185                 190                 195
Ser Cys Pro Ser Arg Asn Leu Glu Leu Phe Lys Ile Gln Arg Ser
                200                 205                 210
Ser Ser Thr Glu Leu Leu Ser His Tyr Asp Ser Asp Ile Lys Glu
                215                 220                 225
Gln Lys Pro Glu His Arg Lys Ser Leu Met Phe Arg Leu His His
                230                 235                 240
Ser Ala Ser Asp Thr Ser Asp Lys Asn Arg Val Lys Ser Arg Leu
                245                 250                 255
Lys Lys Phe Ile Thr Arg Arg Pro Ser Leu Lys Thr Leu Gln Glu
                260                 265                 270
Lys Gly Leu Ile Lys Asp Gln Ile Phe Gly Ser His Leu His Lys
                275                 280                 285
Val Cys Glu Arg Glu Asn Ser Thr Val Pro Trp Phe Val Lys Gln
                290                 295                 300
Cys Ile Glu Ala Val Glu Lys Arg Gly Leu Asp Val Asp Gly Ile
                305                 310                 315
Tyr Arg Val Ser Gly Asn Leu Ala Thr Ile Gln Lys Leu Arg Phe
                320                 325                 330
Ile Val Asn Gln Glu Glu Lys Leu Asn Leu Asp Asp Ser Gln Trp
                335                 340                 345
Glu Asp Ile His Val Val Thr Gly Ala Leu Lys Met Phe Phe Arg
```

```
                    350                     355                     360
Glu Leu Pro Glu Pro Leu Phe Pro Tyr Ser Phe Phe Glu Gln Phe
                    365                     370                     375
Val Glu Ala Ile Lys Lys Gln Asp Asn Asn Thr Arg Ile Glu Ala
                    380                     385                     390
Val Lys Ser Leu Val Gln Lys Leu Pro Pro Pro Asn Arg Asp Thr
                    395                     400                     405
Met Lys Val Leu Phe Gly His Leu Thr Lys Ile Val Ala Lys Ala
                    410                     415                     420
Ser Lys Asn Leu Met Ser Thr Gln Ser Leu Gly Ile Val Phe Gly
                    425                     430                     435
Pro Thr Leu Leu Arg Ala Glu Asn Glu Thr Gly Asn Met Ala Ile
                    440                     445                     450
His Met Val Tyr Gln Asn Gln Ile Ala Glu Leu Met Leu Ser Glu
                    455                     460                     465
Tyr Ser Lys Ile Phe Gly Ser Glu Glu Asp
                    470                     475
```

<210> 20
<211> 368
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 129023CD1

<400> 20

```
Met Ala Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu
  1                 5                      10                      15
Leu Ser His Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser
                    20                      25                      30
Pro Lys Pro Ile Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val
                    35                      40                      45
His Ala Gly Trp Glu Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro
                    50                      55                      60
Tyr Tyr Phe Asn Arg Phe Thr Asn Gln Ser Leu Trp Glu Met Pro
                    65                      70                      75
Val Leu Gly Gln His Asp Val Ile Ser Asp Pro Leu Gly Leu Asn
                    80                      85                      90
Ala Thr Pro Leu Pro Gln Asp Ser Ser Leu Val Glu Thr Pro Pro
                    95                     100                     105
Ala Glu Asn Lys Pro Arg Lys Arg Gln Leu Ser Glu Glu Gln Pro
                    110                     115                     120
Ser Gly Asn Gly Val Lys Lys Pro Lys Ile Glu Ile Pro Val Thr
                    125                     130                     135
Pro Thr Gly Gln Ser Val Pro Ser Ser Pro Ser Ile Pro Gly Thr
                    140                     145                     150
Pro Thr Leu Lys Met Trp Gly Thr Ser Pro Glu Asp Lys Gln Gln
                    155                     160                     165
Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp Asp Leu Asp Ile
                    170                     175                     180
Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser Glu Val Leu
                    185                     190                     195
Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu Ile Leu
                    200                     205                     210
Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu Gly
                    215                     220                     225
Ile Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg
                    230                     235                     240
Lys Val Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys
                    245                     250                     255
Glu Pro Val Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp
                    260                     265                     270
Ile Pro Ile Arg Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys
                    275                     280                     285
Arg Leu Leu Phe Lys Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu
                    290                     295                     300
```

Ser Arg Ser Ala Ser Pro Asp Ser Arg Lys Val Val Lys Trp Asn
305 310 315
Val Glu Asp Thr Leu Ser Trp Leu Arg Glu Gly Pro Leu Ser Leu
320 325 330
Gln Arg Gly Ala Thr Trp Ile Ala Trp Asn Ile Cys Gly Ser Ser
335 340 345
Ala Gly Pro Lys Ser Arg Pro Gln Pro Arg Leu Arg Ala Asn Ala
350 355 360
Ser Ala Val Asp Leu Gln Ile Ser
365

<210> 21
<211> 476
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1358940CD1

<400> 21
Met Phe Arg Leu Asn Ser Leu Ser Ala Leu Ala Glu Leu Ala Val
1 5 10 15
Gly Ser Arg Trp Tyr His Gly Gly Ser Gln Pro Ile Gln Ile Arg
20 25 30
Arg Arg Leu Met Met Val Ala Phe Leu Gly Ala Ser Ala Val Thr
35 40 45
Ala Ser Thr Gly Leu Leu Trp Lys Arg Ala His Ala Glu Ser Pro
50 55 60
Pro Cys Val Asp Asn Leu Lys Ser Asp Ile Gly Asp Lys Gly Lys
65 70 75
Asn Lys Asp Glu Gly Asp Val Cys Asn His Glu Lys Lys Thr Ala
80 85 90
Asp Leu Ala Pro His Pro Glu Glu Lys Lys Lys Lys Arg Ser Gly
95 100 105
Phe Arg Asp Arg Lys Val Met Glu Tyr Glu Asn Arg Ile Arg Ala
110 115 120
Tyr Ser Thr Pro Asp Lys Ile Phe Arg Tyr Phe Ala Thr Leu Lys
125 130 135
Val Ile Ser Glu Pro Gly Glu Ala Glu Val Phe Met Thr Pro Glu
140 145 150
Asp Phe Val Arg Ser Ile Thr Pro Asn Glu Lys Gln Pro Glu His
155 160 165
Leu Gly Leu Asp Gln Tyr Ile Ile Lys Arg Phe Asp Gly Lys Lys
170 175 180
Ile Ser Gln Glu Arg Glu Lys Phe Ala Asp Glu Gly Ser Ile Phe
185 190 195
Tyr Thr Leu Gly Glu Cys Gly Leu Ile Ser Phe Ser Asp Tyr Ile
200 205 210
Phe Leu Thr Thr Val Leu Ser Thr Pro Gln Arg Asn Phe Glu Ile
215 220 225
Ala Phe Lys Met Phe Asp Leu Asn Gly Asp Gly Glu Val Asp Met
230 235 240
Glu Glu Phe Glu Gln Val Gln Ser Ile Ile Arg Ser Gln Thr Ser
245 250 255
Met Gly Met Arg His Arg Asp Arg Pro Thr Thr Gly Asn Thr Leu
260 265 270
Lys Ser Gly Leu Cys Ser Ala Leu Thr Thr Tyr Phe Phe Gly Ala
275 280 285
Asp Leu Lys Gly Lys Leu Thr Ile Lys Asn Phe Leu Glu Phe Gln
290 295 300
Arg Lys Leu Gln His Asp Val Leu Lys Leu Glu Phe Glu Arg His
305 310 315
Asp Pro Val Asp Gly Arg Ile Thr Glu Arg Gln Phe Gly Gly Met
320 325 330
Leu Leu Ala Tyr Ser Gly Val Gln Ser Lys Lys Leu Thr Ala Met
335 340 345
Gln Arg Gln Leu Lys Lys His Phe Lys Glu Gly Lys Gly Leu Thr

```
                    350                           355                           360
Phe Gln Glu Val Glu Asn Phe Phe Thr Phe Leu Lys Asn Ile Asn
                    365                           370                           375
Asp Val Asp Thr Ala Leu Ser Phe Tyr His Met Ala Gly Ala Ser
                    380                           385                           390
Leu Asp Lys Val Thr Met Gln Gln Val Ala Arg Thr Val Ala Lys
                    395                           400                           405
Val Glu Leu Ser Asp His Val Cys Asp Val Val Phe Ala Leu Phe
                    410                           415                           420
Asp Cys Asp Gly Asn Gly Glu Leu Ser Asn Lys Glu Phe Val Ser
                    425                           430                           435
Ile Met Lys Gln Arg Leu Met Arg Gly Leu Glu Lys Pro Lys Asp
                    440                           445                           450
Met Gly Phe Thr Arg Leu Met Gln Ala Met Trp Lys Cys Ala Gln
                    455                           460                           465
Glu Thr Ala Trp Asp Phe Ala Leu Pro Lys Gln
                    470                           475
<210> 22
<211> 171
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1682320CD1

<400> 22
Met Glu Lys Arg Leu Gln Glu Ala Gln Leu Tyr Lys Glu Glu Gly
  1                 5                            10                            15
Asn Gln Arg Tyr Arg Glu Gly Lys Tyr Arg Asp Ala Val Ser Arg
                    20                           25                            30
Tyr His Arg Ala Leu Leu Gln Leu Arg Gly Leu Asp Pro Ser Leu
                    35                           40                            45
Pro Ser Pro Leu Pro Asn Leu Gly Pro Gln Gly Pro Ala Leu Thr
                    50                           55                            60
Pro Glu Gln Glu Asn Ile Leu His Thr Thr Gln Thr Asp Cys Tyr
                    65                           70                            75
Asn Asn Leu Ala Ala Cys Leu Leu Gln Met Glu Pro Val Asn Tyr
                    80                           85                            90
Glu Arg Val Arg Glu Tyr Ser Gln Lys Val Leu Glu Arg Gln Pro
                    95                           100                           105
Asp Asn Ala Lys Ala Leu Tyr Arg Ala Gly Val Ala Phe Phe His
                    110                           115                           120
Leu Gln Asp Tyr Asp Gln Ala Arg His Tyr Leu Leu Ala Ala Val
                    125                           130                           135
Asn Arg Gln Pro Lys Asp Ala Asn Val Arg Arg Tyr Leu Gln Leu
                    140                           145                           150
Thr Gln Ser Glu Leu Ser Ser Tyr His Arg Lys Glu Lys Gln Leu
                    155                           160                           165
Tyr Leu Gly Met Phe Gly
                    170
<210> 23
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1728263CD1

<400> 23
Met Phe Phe Ser Glu Ala Arg Ala Arg Ser Arg Thr Trp Glu Ala
  1                 5                            10                            15
Ser Pro Ser Glu His Arg Lys Trp Val Glu Val Phe Lys Ala Cys
                    20                           25                            30
Asp Glu Asp His Lys Gly Tyr Leu Ser Arg Glu Asp Phe Lys Thr
                    35                           40                            45
```

```
Ala Val Val Met Leu Phe Gly Tyr Lys Pro Ser Lys Ile Glu Val
                50                  55                  60
Asp Ser Val Met Ser Ser Ile Asn Pro Asn Thr Ser Gly Ile Leu
                65                  70                  75
Leu Glu Gly Phe Leu Asn Ile Val Arg Lys Lys Lys Glu Ala Gln
                80                  85                  90
Arg Tyr Arg Asn Glu Val Arg His Ile Phe Thr Ala Phe Asp Thr
                95                  100                 105
Tyr Tyr Arg Gly Phe Leu Thr Leu Glu Asp Phe Lys Lys Ala Phe
                110                 115                 120
Arg Gln Val Ala Pro Lys Leu Pro Glu Arg Thr Val Leu Glu Val
                125                 130                 135
Phe Arg Glu Val Asp Arg Asp Ser Asp Gly His Val Ser Phe Arg
                140                 145                 150
Asp Phe Glu Tyr Ala Leu Asn Tyr Gly Gln Lys Glu Ala
                155                 160
<210> 24
<211> 354
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1867626CD1

<400> 24
Met Gly Glu Gln Pro Ile Phe Ser Thr Arg Ala His Val Phe Gln
 1               5                   10                  15
Ile Asp Pro Asn Thr Lys Lys Asn Trp Val Pro Thr Ser Lys His
                20                  25                  30
Ala Val Thr Val Ser Tyr Phe Tyr Asp Ser Thr Arg Asn Val Tyr
                35                  40                  45
Arg Ile Ile Ser Leu Asp Gly Ser Lys Ala Ile Ile Asn Ser Thr
                50                  55                  60
Ile Thr Pro Asn Met Thr Phe Thr Lys Thr Ser Gln Arg Phe Gly
                65                  70                  75
Gln Trp Ala Asp Ser Arg Ala Asn Thr Val Tyr Gly Leu Gly Phe
                80                  85                  90
Ser Ser Glu His His Leu Ser Lys Phe Ala Glu Lys Phe Gln Glu
                95                  100                 105
Phe Lys Glu Ala Ala Arg Leu Ala Lys Glu Lys Ser Gln Glu Lys
                110                 115                 120
Met Glu Leu Thr Ser Thr Pro Ser Gln Glu Ser Ala Gly Gly Asp
                125                 130                 135
Leu Gln Ser Pro Leu Thr Pro Glu Ser Ile Asn Gly Thr Asp Asp
                140                 145                 150
Glu Arg Thr Pro Asp Val Thr Gln Asn Ser Glu Pro Arg Ala Glu
                155                 160                 165
Pro Thr Gln Asn Ala Leu Pro Phe Ser His Ser Ser Ala Ile Ser
                170                 175                 180
Lys His Trp Glu Ala Glu Leu Ala Thr Leu Lys Gly Asn Asn Ala
                185                 190                 195
Lys Leu Thr Ala Ala Leu Leu Glu Ser Thr Ala Asn Val Lys Gln
                200                 205                 210
Trp Lys Gln Gln Leu Ala Ala Tyr Gln Glu Glu Ala Glu Arg Leu
                215                 220                 225
His Lys Arg Val Thr Glu Leu Glu Cys Val Ser Ser Gln Ala Asn
                230                 235                 240
Ala Val His Thr His Lys Thr Glu Leu Asn Gln Thr Ile Gln Glu
                245                 250                 255
Leu Glu Glu Thr Leu Lys Leu Lys Glu Glu Glu Ile Glu Arg Leu
                260                 265                 270
Lys Gln Glu Ile Asp Asn Ala Arg Glu Leu Gln Glu Gln Arg Asp
                275                 280                 285
Ser Leu Thr Gln Lys Leu Gln Glu Val Glu Ile Arg Asn Lys Asp
                290                 295                 300
Leu Glu Gly Gln Leu Ser Asp Leu Glu Gln Arg Leu Glu Lys Ser
```

305 310 315
Gln Asn Glu Gln Glu Ala Phe Arg Asn Asn Leu Lys Thr Leu Leu
320 325 330
Glu Ile Leu Asp Gly Lys Ile Phe Glu Leu Thr Glu Leu Arg Asp
335 340 345
Asn Leu Ala Lys Leu Leu Glu Cys Ser
350

<210> 25
<211> 365
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1990126CD1

<400> 25
Met Asn Ile Met Asp Phe Asn Val Lys Lys Leu Ala Ala Asp Ala
1 5 10 15
Gly Thr Phe Leu Ser Arg Ala Val Gln Phe Thr Glu Glu Lys Leu
20 25 30
Gly Gln Ala Glu Lys Thr Glu Leu Asp Ala His Leu Glu Asn Leu
35 40 45
Leu Ser Lys Ala Glu Cys Thr Lys Ile Trp Thr Glu Lys Ile Met
50 55 60
Lys Gln Thr Glu Val Leu Leu Gln Pro Asn Pro Asn Ala Arg Ile
65 70 75
Glu Glu Phe Val Tyr Glu Lys Leu Asp Arg Lys Ala Pro Ser Arg
80 85 90
Ile Asn Asn Pro Glu Leu Leu Gly Gln Tyr Met Ile Asp Ala Gly
95 100 105
Thr Glu Phe Gly Pro Gly Thr Ala Tyr Gly Asn Ala Leu Ile Lys
110 115 120
Cys Gly Glu Thr Gln Lys Arg Ile Gly Thr Ala Asp Arg Glu Leu
125 130 135
Ile Gln Thr Ser Ala Leu Asn Phe Leu Thr Pro Leu Arg Asn Phe
140 145 150
Ile Glu Gly Asp Tyr Lys Thr Ile Ala Lys Glu Arg Lys Leu Leu
155 160 165
Gln Asn Lys Arg Leu Asp Leu Asp Ala Ala Lys Thr Arg Leu Lys
170 175 180
Lys Ala Lys Ala Ala Glu Thr Arg Asn Ser Ser Glu Gln Glu Leu
185 190 195
Arg Ile Thr Gln Ser Glu Phe Asp Arg Gln Ala Glu Ile Thr Arg
200 205 210
Leu Leu Leu Glu Gly Ile Ser Ser Thr His Ala His His Leu Arg
215 220 225
Cys Leu Asn Asp Phe Val Glu Ala Gln Met Thr Tyr Tyr Ala Gln
230 235 240
Cys Tyr Gln Tyr Met Leu Asp Leu Gln Lys Gln Leu Gly Ser Phe
245 250 255
Pro Ser Asn Tyr Leu Ser Asn Asn Asn Gln Thr Ser Val Thr Pro
260 265 270
Val Pro Ser Val Leu Pro Asn Ala Ile Gly Ser Ser Ala Met Ala
275 280 285
Ser Thr Ser Gly Leu Val Ile Thr Ser Pro Ser Asn Leu Ser Asp
290 295 300
Leu Lys Glu Cys Ser Gly Ser Arg Lys Ala Arg Val Leu Tyr Asp
305 310 315
Tyr Asp Ala Ala Asn Ser Thr Glu Leu Ser Leu Leu Ala Asp Glu
320 325 330
Val Ile Thr Val Phe Ser Val Val Gly Met Asp Ser Asp Trp Leu
335 340 345
Met Gly Glu Arg Gly Asn Gln Lys Gly Lys Val Pro Ile Thr Tyr
350 355 360
Leu Glu Leu Leu Asn
365

<210> 26
<211> 274
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2104180CD1

<400> 26
Met Ala Thr Thr Val Ser Thr Gln Arg Gly Pro Val Tyr Ile Gly
1 5 10 15
Glu Leu Pro Gln Asp Phe Leu Arg Ile Thr Pro Thr Gln Gln Gln
20 25 30
Arg Gln Val Gln Leu Asp Ala Gln Ala Ala Gln Gln Leu Gln Tyr
35 40 45
Gly Gly Ala Val Gly Thr Val Gly Arg Leu Asn Ile Thr Val Val
50 55 60
Gln Ala Lys Leu Ala Lys Asn Tyr Gly Met Thr Arg Met Asp Pro
65 70 75
Tyr Cys Arg Leu Arg Leu Gly Tyr Ala Val Tyr Glu Thr Pro Thr
80 85 90
Ala His Asn Gly Ala Lys Asn Pro Arg Trp Asn Lys Val Ile His
95 100 105
Cys Thr Val Pro Pro Gly Val Asp Ser Phe Tyr Leu Glu Ile Phe
110 115 120
Asp Glu Arg Ala Phe Ser Met Asp Asp Arg Ile Ala Trp Thr His
125 130 135
Ile Thr Ile Pro Glu Ser Leu Arg Gln Gly Lys Val Glu Asp Lys
140 145 150
Trp Tyr Ser Leu Ser Gly Arg Gln Gly Asp Asp Lys Glu Gly Met
155 160 165
Ile Asn Leu Val Met Ser Tyr Ala Leu Leu Pro Ala Ala Met Val
170 175 180
Met Pro Pro Gln Pro Val Val Leu Met Pro Thr Val Tyr Gln Gln
185 190 195
Gly Val Gly Tyr Val Pro Ile Thr Gly Met Pro Ala Val Cys Ser
200 205 210
Pro Gly Met Val Pro Val Ala Leu Pro Pro Ala Ala Val Asn Ala
215 220 225
Gln Pro Arg Cys Ser Glu Glu Asp Leu Lys Ala Ile Gln Asp Met
230 235 240
Phe Pro Asn Met Asp Gln Glu Val Ile Arg Ser Val Leu Glu Ala
245 250 255
Gln Arg Gly Asn Lys Asp Ala Ala Ile Asn Ser Leu Leu Gln Met
260 265 270
Gly Glu Glu Pro

<210> 27
<211> 129
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2122241CD1

<400> 27
Met Arg Arg Arg Gly Glu Ile Asp Met Ala Thr Glu Gly Asp Val
1 5 10 15
Glu Leu Glu Leu Glu Thr Glu Thr Ser Gly Pro Glu Arg Pro Pro
20 25 30
Glu Lys Pro Arg Lys His Asp Ser Gly Ala Ala Asp Leu Glu Arg
35 40 45
Val Thr Asp Tyr Ala Glu Glu Lys Glu Ile Gln Ser Ser Asn Leu
50 55 60
Glu Thr Ala Met Ser Val Ile Gly Asp Arg Arg Ser Arg Glu Gln

65 70 75
Lys Ala Lys Gln Glu Arg Glu Lys Glu Leu Ala Lys Val Thr Ile
80 85 90
Lys Lys Glu Asp Leu Glu Leu Ile Met Thr Glu Met Glu Ile Ser
95 100 105
Arg Ala Ala Ala Glu Arg Ser Leu Arg Glu His Met Gly Asn Val
110 115 120
Val Glu Ala Leu Ile Ala Leu Thr Asn
125

<210> 28
<211> 626
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2580428CD1

<400> 28
Met Gln Arg Ala Asp Ser Glu Gln Pro Ser Lys Arg Pro Arg Cys
1 5 10 15
Asp Asp Ser Pro Arg Thr Pro Ser Asn Thr Pro Ser Ala Glu Ala
20 25 30
Asp Trp Ser Pro Gly Leu Glu Leu His Pro Asp Tyr Lys Thr Trp
35 40 45
Gly Pro Glu Gln Val Cys Ser Phe Leu Arg Arg Gly Gly Phe Glu
50 55 60
Glu Pro Val Leu Leu Lys Asn Ile Arg Glu Asn Glu Ile Thr Gly
65 70 75
Ala Leu Leu Pro Cys Leu Asp Glu Ser Arg Phe Glu Asn Leu Gly
80 85 90
Val Ser Ser Leu Gly Glu Arg Lys Lys Leu Leu Ser Tyr Ile Gln
95 100 105
Arg Leu Val Gln Ile His Val Asp Thr Met Lys Val Ile Asn Asp
110 115 120
Pro Ile His Gly His Ile Glu Leu His Pro Leu Leu Val Arg Ile
125 130 135
Ile Asp Thr Pro Gln Phe Gln Arg Leu Arg Tyr Ile Lys Gln Leu
140 145 150
Gly Gly Gly Tyr Tyr Val Phe Pro Gly Ala Ser His Asn Arg Phe
155 160 165
Glu His Ser Leu Gly Val Gly Tyr Leu Ala Gly Cys Leu Val His
170 175 180
Ala Leu Gly Glu Lys Gln Pro Glu Leu Gln Ile Ser Glu Arg Asp
185 190 195
Val Leu Cys Val Gln Ile Ala Gly Leu Cys His Asp Leu Gly His
200 205 210
Gly Pro Phe Ser His Met Phe Asp Gly Arg Phe Ile Pro Leu Ala
215 220 225
Arg Pro Glu Val Lys Trp Thr His Glu Gln Gly Ser Val Met Met
230 235 240
Phe Glu His Leu Ile Asn Ser Asn Gly Ile Lys Pro Val Met Glu
245 250 255
Gln Tyr Gly Leu Ile Pro Glu Glu Asp Ile Cys Phe Ile Lys Glu
260 265 270
Gln Ile Val Gly Pro Leu Glu Ser Pro Val Glu Asp Ser Leu Trp
275 280 285
Pro Tyr Lys Gly Arg Pro Glu Asn Lys Ser Phe Leu Tyr Glu Ile
290 295 300
Val Ser Asn Lys Arg Asn Gly Ile Asp Val Asp Lys Trp Asp Tyr
305 310 315
Phe Ala Arg Asp Cys His His Leu Gly Ile Gln Asn Asn Phe Asp
320 325 330
Tyr Lys Arg Phe Ile Lys Phe Ala Arg Val Cys Glu Val Asp Asn
335 340 345
Glu Leu Arg Ile Cys Ala Arg Asp Lys Glu Val Gly Asn Leu Tyr
350 355 360

Asp Met Phe His Thr Arg Asn Ser Leu His Arg Arg Ala Tyr Gln
365 370 375
His Lys Val Gly Asn Ile Ile Asp Thr Met Ile Thr Asp Ala Phe
380 385 390
Leu Lys Ala Asp Asp Tyr Ile Glu Ile Thr Gly Ala Gly Gly Lys
395 400 405
Lys Tyr Arg Ile Ser Thr Ala Ile Asp Asp Met Glu Ala Tyr Thr
410 415 420
Lys Leu Thr Asp Asn Ile Phe Leu Glu Ile Leu Tyr Ser Thr Asp
425 430 435
Pro Lys Leu Lys Asp Ala Arg Glu Ile Leu Lys Gln Ile Glu Tyr
440 445 450
Arg Asn Leu Phe Lys Tyr Val Gly Glu Thr Gln Pro Thr Gly Gln
455 460 465
Ile Lys Ile Lys Arg Glu Asp Tyr Glu Ser Leu Pro Lys Glu Val
470 475 480
Ala Ser Ala Lys Pro Lys Val Leu Leu Asp Val Lys Leu Lys Ala
485 490 495
Glu Asp Phe Ile Val Asp Val Ile Asn Met Asp Tyr Gly Met Gln
500 505 510
Glu Lys Asn Pro Ile Asp His Val Ser Phe Tyr Cys Lys Thr Ala
515 520 525
Pro Asn Arg Ala Ile Arg Ile Thr Lys Asn Gln Val Ser Gln Leu
530 535 540
Leu Pro Glu Lys Phe Ala Glu Gln Leu Ile Arg Val Tyr Cys Lys
545 550 555
Lys Val Asp Arg Lys Ser Leu Tyr Ala Ala Arg Gln Tyr Phe Val
560 565 570
Gln Trp Cys Ala Asp Arg Asn Phe Thr Lys Pro Gln Asp Gly Asp
575 580 585
Val Ile Ala Pro Leu Ile Thr Pro Gln Lys Lys Glu Trp Asn Asp
590 595 600
Ser Thr Ser Val Gln Asn Pro Thr Arg Leu Arg Glu Ala Ser Lys
605 610 615
Ser Arg Val Gln Leu Phe Lys Asp Asp Pro Met
620 625

<210> 29
<211> 157
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3397189CD1

<400> 29
Met Ala Pro Lys Lys Leu Ser Cys Leu Arg Ser Leu Leu Leu Pro
1 5 10 15
Leu Ser Leu Thr Leu Leu Leu Pro Gln Ala Asp Thr Arg Ser Phe
20 25 30
Val Val Asp Arg Gly His Asp Arg Phe Leu Leu Asp Gly Ala Pro
35 40 45
Phe Arg Tyr Val Ser Gly Ser Leu His Tyr Phe Arg Val Pro Arg
50 55 60
Val Leu Trp Ala Asp Arg Leu Leu Lys Met Arg Trp Ser Gly Leu
65 70 75
Asn Ala Ile Gln Phe Tyr Val Pro Trp Asn Tyr His Glu Pro Gln
80 85 90
Pro Gly Val Tyr Asn Phe Asn Gly Ser Arg Asp Leu Ile Ala Phe
95 100 105
Leu Asn Glu Ala Ala Leu Ala Asn Leu Leu Val Ile Leu Arg Pro
110 115 120
Gly Pro Tyr Ile Cys Ala Glu Trp Glu Met Gly Gly Leu Pro Ser
125 130 135
Trp Leu Leu Arg Lys Pro Glu Ile His Leu Arg Thr Ser Asp Pro
140 145 150
Gly Glu Leu Arg Gln Arg Ile

155
<210> 30
<211> 383
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4881249CD1

<400> 30
Met Leu Ser Arg Lys Lys Thr Lys Asn Glu Val Ser Lys Pro Ala
1 5 10 15
Glu Val Gln Gly Lys Tyr Val Lys Lys Glu Thr Ser Pro Leu Leu
20 25 30
Arg Asn Leu Met Pro Ser Phe Ile Arg His Gly Pro Thr Ile Pro
35 40 45
Arg Arg Thr Asp Ile Cys Leu Pro Asp Ser Ser Pro Asn Ala Phe
50 55 60
Ser Thr Ser Gly Asp Val Val Ser Arg Asn Gln Ser Phe Leu Arg
65 70 75
Thr Pro Ile Gln Arg Thr Pro His Glu Ile Met Arg Arg Glu Ser
80 85 90
Asn Arg Leu Ser Ala Pro Ser Tyr Leu Ala Arg Ser Leu Ala Asp
95 100 105
Val Pro Arg Glu Tyr Gly Ser Ser Gln Ser Phe Val Thr Glu Val
110 115 120
Ser Phe Ala Val Glu Asn Gly Asp Ser Gly Ser Arg Tyr Tyr Tyr
125 130 135
Ser Asp Asn Phe Phe Asp Gly Gln Arg Lys Arg Pro Leu Gly Asp
140 145 150
Arg Ala His Glu Asp Tyr Arg Tyr Tyr Glu Tyr Asn His Asp Leu
155 160 165
Phe Gln Arg Met Pro Gln Asn Gln Gly Arg His Ala Ser Gly Ile
170 175 180
Gly Arg Val Ala Ala Thr Ser Leu Gly Asn Leu Thr Asn His Gly
185 190 195
Ser Glu Asp Leu Pro Leu Pro Pro Gly Trp Ser Val Asp Trp Thr
200 205 210
Met Arg Gly Arg Lys Tyr Tyr Ile Asp His Asn Thr Asn Thr Thr
215 220 225
His Trp Ser His Pro Leu Glu Arg Glu Gly Leu Pro Pro Gly Trp
230 235 240
Glu Arg Val Glu Ser Ser Glu Phe Gly Thr Tyr Tyr Val Asp His
245 250 255
Thr Asn Lys Lys Ala Gln Tyr Arg His Pro Cys Ala Pro Ser Val
260 265 270
Pro Arg Tyr Asp Gln Pro Pro Pro Val Thr Tyr Gln Pro Gln Gln
275 280 285
Thr Glu Arg Asn Gln Ser Leu Leu Val Pro Ala Asn Pro Tyr His
290 295 300
Thr Ala Glu Ile Pro Asp Trp Leu Gln Val Tyr Ala Arg Ala Pro
305 310 315
Val Lys Tyr Asp His Ile Leu Lys Trp Glu Leu Phe Gln Leu Ala
320 325 330
Asp Leu Asp Thr Tyr Gln Gly Met Leu Lys Leu Leu Phe Met Lys
335 340 345
Glu Leu Glu Gln Ile Val Lys Met Tyr Glu Ala Tyr Arg Gln Ala
350 355 360
Leu Leu Thr Glu Leu Glu Asn Arg Lys Gln Arg Gln Gln Trp Tyr
365 370 375
Ala Gln Gln His Gly Lys Asn Phe
380
<210> 31
<211> 478
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 431871CD1

<400> 31
Met Asp Thr Ser Asp Leu Phe Ala Ser Cys Arg Lys Gly Asp Val
1 5 10 15
Gly Arg Val Arg Tyr Leu Leu Glu Gln Arg Asp Val Glu Val Asn
20 25 30
Val Arg Asp Lys Trp Asp Ser Thr Pro Leu Tyr Tyr Ala Cys Leu
35 40 45
Cys Gly His Glu Glu Leu Val Leu Tyr Leu Leu Ala Asn Gly Ala
50 55 60
Arg Cys Glu Ala Asn Thr Phe Asp Gly Glu Arg Cys Leu Tyr Gly
65 70 75
Ala Leu Ser Asp Pro Ile Arg Arg Ala Leu Arg Asp Tyr Lys Gln
80 85 90
Val Thr Ala Ser Cys Arg Arg Arg Asp Tyr Tyr Asp Asp Phe Leu
95 100 105
Gln Arg Leu Leu Glu Gln Gly Ile His Ser Asp Val Val Phe Val
110 115 120
Val His Gly Lys Pro Phe Arg Val His Arg Cys Val Leu Gly Ala
125 130 135
Arg Ser Ala Tyr Phe Ala Asn Met Leu Asp Thr Lys Trp Lys Gly
140 145 150
Lys Ser Val Val Val Leu Arg His Pro Leu Ile Asn Pro Val Ala
155 160 165
Phe Gly Ala Leu Leu Gln Tyr Leu Tyr Thr Gly Arg Leu Asp Ile
170 175 180
Gly Val Glu His Val Ser Asp Cys Glu Arg Leu Ala Lys Gln Cys
185 190 195
Gln Leu Trp Asp Leu Leu Ser Asp Leu Glu Ala Lys Cys Glu Lys
200 205 210
Val Ser Glu Phe Val Ala Ser Lys Pro Gly Thr Cys Val Lys Val
215 220 225
Leu Thr Ile Glu Pro Pro Pro Ala Asp Pro Arg Leu Arg Glu Asp
230 235 240
Met Ala Leu Leu Ala Asp Cys Ala Leu Pro Pro Glu Leu Arg Gly
245 250 255
Asp Leu Trp Glu Leu Pro Phe Pro Cys Pro Asp Gly Phe Asn Ser
260 265 270
Cys Pro Asp Ile Cys Phe Arg Val Ala Gly Cys Ser Phe Leu Cys
275 280 285
His Lys Ala Phe Phe Cys Gly Arg Ser Asp Tyr Phe Arg Ala Leu
290 295 300
Leu Asp Asp His Phe Arg Glu Ser Glu Glu Pro Ala Thr Ser Gly
305 310 315
Gly Pro Pro Ala Val Thr Leu His Gly Ile Ser Pro Asp Val Phe
320 325 330
Thr His Val Leu Tyr Tyr Met Tyr Ser Asp His Thr Glu Leu Ser
335 340 345
Pro Glu Ala Ala Tyr Asp Val Leu Ser Val Ala Asp Met Tyr Leu
350 355 360
Leu Pro Gly Leu Lys Arg Leu Cys Gly Arg Ser Leu Ala Gln Met
365 370 375
Leu Asp Glu Asp Thr Val Val Gly Val Trp Arg Val Ala Lys Leu
380 385 390
Phe Arg Leu Ala Arg Leu Glu Asp Gln Cys Thr Glu Tyr Met Ala
395 400 405
Lys Val Ile Glu Lys Leu Val Glu Arg Glu Asp Phe Val Glu Ala
410 415 420
Val Lys Glu Glu Ala Ala Ala Val Ala Ala Arg Gln Glu Thr Asp
425 430 435
Ser Ile Pro Leu Val Asp Asp Ile Arg Phe His Val Ala Ser Thr
440 445 450
Val Gln Thr Tyr Ser Ala Ile Glu Glu Ala Gln Gln Arg Leu Arg
455 460 465

Ala Leu Glu Asp Leu Leu Val Ser Ile Gly Leu Asp Cys
470 475
<210> 32
<211> 275
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 526155CD1

<400> 32
Met Ser Ala Glu Val Lys Val Thr Gly Gln Asn Gln Glu Gln Phe
1 5 10 15
Leu Leu Leu Ala Lys Ser Ala Lys Gly Ala Ala Leu Ala Thr Leu
20 25 30
Ile His Gln Val Leu Glu Ala Pro Gly Val Tyr Val Phe Gly Glu
35 40 45
Leu Leu Asp Met Pro Asn Val Arg Glu Leu Ala Glu Ser Asp Phe
50 55 60
Ala Ser Thr Phe Arg Leu Leu Thr Val Phe Ala Tyr Gly Thr Tyr
65 70 75
Ala Asp Tyr Leu Ala Glu Ala Arg Asn Leu Pro Pro Leu Thr Glu
80 85 90
Ala Gln Lys Asn Lys Leu Arg His Leu Ser Val Val Thr Leu Ala
95 100 105
Ala Lys Val Lys Cys Ile Pro Tyr Ala Val Leu Leu Glu Ala Leu
110 115 120
Ala Leu Arg Asn Val Arg Gln Leu Glu Asp Leu Val Ile Glu Ala
125 130 135
Val Tyr Ala Asp Val Leu Arg Gly Ser Leu Asp Gln Arg Asn Gln
140 145 150
Arg Leu Glu Val Asp Tyr Ser Ile Gly Arg Asp Ile Gln Arg Gln
155 160 165
Asp Leu Ser Ala Ile Ala Arg Thr Leu Gln Glu Trp Cys Val Gly
170 175 180
Cys Glu Val Val Leu Ser Gly Ile Glu Glu Gln Val Ser Arg Ala
185 190 195
Asn Gln His Lys Glu Gln Gln Leu Gly Leu Lys Gln Gln Ile Glu
200 205 210
Ser Glu Val Ala Asn Leu Lys Lys Thr Ile Lys Val Thr Thr Ala
215 220 225
Ala Ala Ala Ala Ala Thr Ser Gln Asp Pro Glu Gln His Leu Thr
230 235 240
Glu Leu Arg Glu Pro Ala Pro Gly Thr Asn Gln Arg Gln Pro Ser
245 250 255
Lys Lys Ala Ser Lys Gly Lys Gly Leu Arg Gly Ser Ala Lys Ile
260 265 270
Trp Ser Lys Ser Asn
275
<210> 33
<211> 217
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 676234CD1

<400> 33
Met Ala Ser Thr Gly Leu Glu Leu Leu Gly Met Thr Leu Ala Val
1 5 10 15
Leu Gly Trp Leu Gly Thr Leu Val Ser Cys Ala Leu Pro Leu Trp
20 25 30
Lys Val Thr Ala Phe Ile Gly Asn Ser Ile Val Val Ala Gln Val

```
                    35                       40                       45
Val Trp Glu Gly Leu Trp Met Ser Cys Val Val Gln Ser Thr Gly
                    50                       55                       60
Gln Met Gln Cys Lys Val Tyr Asp Ser Leu Leu Ala Leu Pro Gln
                    65                       70                       75
Asp Leu Gln Ala Ala Arg Ala Leu Cys Val Ile Ala Leu Leu Leu
                    80                       85                       90
Ala Leu Leu Gly Leu Leu Val Ala Ile Thr Gly Ala Gln Cys Thr
                    95                      100                      105
Thr Cys Val Glu Asp Glu Gly Ala Lys Ala Arg Ile Val Leu Thr
                   110                      115                      120
Ala Gly Val Ile Leu Leu Leu Ala Gly Ile Leu Val Leu Ile Pro
                   125                      130                      135
Val Cys Trp Thr Ala His Ala Ile Ile Gln Asp Phe Tyr Asn Pro
                   140                      145                      150
Leu Val Ala Glu Ala Leu Lys Arg Glu Leu Gly Ala Ser Leu Tyr
                   155                      160                      165
Leu Gly Trp Ala Ala Ala Ala Leu Leu Met Leu Gly Gly Gly Leu
                   170                      175                      180
Leu Cys Cys Thr Cys Pro Pro Pro Gln Val Glu Arg Pro Arg Gly
                   185                      190                      195
Pro Arg Leu Gly Tyr Ser Ile Pro Ser Arg Ser Gly Ala Ser Gly
                   200                      205                      210
Leu Asp Lys Arg Asp Tyr Val
                   215
```

<210> 34
<211> 74
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 720145CD1

<400> 34

```
Met Asp Asp Tyr Thr Ser Ala Ile Glu Val Gln Pro Asn Phe Glu
  1                 5                       10                       15
Val Pro Tyr Tyr Asn Arg Gly Leu Ile Leu Tyr Arg Leu Gly Tyr
                    20                       25                       30
Phe Asp Asp Ala Leu Glu Asp Phe Lys Lys Val Leu Asp Leu Asn
                    35                       40                       45
Pro Gly Phe Gln Asp Ala Thr Leu Ser Leu Lys Gln Thr Ile Leu
                    50                       55                       60
Asp Lys Glu Glu Lys Gln Arg Arg Asn Val Ala Lys Asn Tyr
                    65                       70
```

<210> 35
<211> 367
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1001951CD1

<400> 35

```
Met Val Gln Gln Phe Leu Arg Gln Ala Gln Arg Gly Thr Glu Glu
  1                 5                       10                       15
Lys Glu Arg Glu Gly Ala Leu Val Ser Leu Arg Arg Gly Leu Gln
                    20                       25                       30
His Pro Glu Thr Gln Gln Thr Phe Ile Arg Ser Cys Val Cys Ile
                    35                       40                       45
His Trp Val Thr Leu Ile Val Glu Ser Glu Ala Val Arg Arg Gln
                    50                       55                       60
Leu Leu Pro Gln Gly Ile Val Pro Ala Leu Ala Ala Cys Ile Gln
                    65                       70                       75
Ser Pro His Val Ala Val Leu Glu Ala Leu Gly Tyr Ala Leu Ser
                    80                       85                       90
```

```
Gln Leu Leu Gln Ala Glu Glu Ala Pro Glu Lys Ile Ile Pro Ser
                95                  100                 105
Ile Leu Ala Ser Thr Leu Pro Gln His Met Leu Gln Met Leu Gln
                110                 115                 120
Pro Gly Pro Lys Leu Asn Pro Gly Val Ala Val Glu Phe Ala Trp
                125                 130                 135
Cys Leu His Tyr Ile Ile Cys Ser Gln Val Ser Asn Pro Leu Leu
                140                 145                 150
Ile Gly His Gly Ala Leu Ser Thr Leu Gly Leu Leu Leu Leu Asp
                155                 160                 165
Leu Ala Gly Ala Val Gln Lys Thr Glu Asp Ala Gly Leu Glu Leu
                170                 175                 180
Leu Ala Cys Pro Val Leu Arg Cys Leu Ser Asn Leu Leu Thr Glu
                185                 190                 195
Ala Ala Val Glu Thr Val Gly Gly Gln Met Gln Leu Arg Asp Glu
                200                 205                 210
Arg Val Val Ala Ala Leu Phe Ile Leu Leu Gln Phe Phe Phe Gln
                215                 220                 225
Lys Gln Pro Ser Leu Leu Pro Glu Gly Leu Trp Leu Leu Asn Asn
                230                 235                 240
Leu Thr Ala Asn Ser Pro Ser Phe Cys Thr Ser Leu Leu Ser Leu
                245                 250                 255
Asp Leu Ile Glu Pro Leu Leu Gln Leu Leu Pro Val Ser Asn Val
                260                 265                 270
Val Ser Val Met Val Leu Thr Val Leu Cys Asn Val Ala Glu Lys
                275                 280                 285
Gly Pro Ala Tyr Cys Gln Arg Leu Trp Pro Gly Pro Leu Leu Pro
                290                 295                 300
Ala Leu Leu His Thr Leu Ala Phe Ser Asp Thr Glu Val Val Gly
                305                 310                 315
Gln Ser Leu Glu Leu Leu His Leu Leu Phe Leu Tyr Gln Pro Glu
                320                 325                 330
Ala Val Gln Val Phe Leu Gln Gln Ser Gly Leu Gln Ala Trp Lys
                335                 340                 345
Arg His Gln Glu Glu Ala Gln Leu Gln Asp Arg Val Tyr Ala Leu
                350                 355                 360
Gln Gln Thr Ala Leu Gln Gly
                365
```

<210> 36
<211> 1113
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1243349CD1

<400> 36

```
Met Ile Ala Val Ser Phe Lys Cys Arg Cys Gln Ile Leu Arg Arg
 1               5                  10                  15
Leu Thr Lys Asp Glu Ser Pro Tyr Thr Lys Ser Ala Ser Gln Thr
                20                  25                  30
Lys Pro Pro Asp Gly Ala Leu Ala Val Arg Arg Gln Ser Ile Pro
                35                  40                  45
Glu Glu Phe Lys Gly Ser Thr Val Val Glu Leu Met Lys Lys Glu
                50                  55                  60
Gly Thr Thr Leu Gly Leu Thr Val Ser Gly Gly Ile Asp Lys Asp
                65                  70                  75
Gly Lys Pro Arg Val Ser Asn Leu Arg Gln Gly Gly Ile Ala Ala
                80                  85                  90
Arg Ser Asp Gln Leu Asp Val Gly Asp Tyr Ile Lys Ala Val Asn
                95                  100                 105
Gly Ile Asn Leu Ala Lys Phe Arg His Asp Glu Ile Ile Ser Leu
                110                 115                 120
Leu Lys Asn Val Gly Glu Arg Val Val Leu Glu Val Glu Tyr Glu
                125                 130                 135
Leu Pro Pro Val Ser Val Gln Gly Ser Ser Val Ile Phe Arg Thr
```

```
                    140                           145                           150
Val Glu Val Thr Leu His Lys Glu Gly Asn Thr Phe Gly Phe Val
                    155                           160                           165
Ile Arg Gly Gly Ala His Asp Asp Arg Asn Lys Ser Arg Pro Val
                    170                           175                           180
Val Ile Thr Cys Val Arg Pro Gly Gly Pro Ala Asp Arg Glu Gly
                    185                           190                           195
Thr Ile Lys Pro Gly Asp Arg Leu Leu Ser Val Asp Gly Ile Arg
                    200                           205                           210
Leu Leu Gly Thr Thr His Ala Glu Ala Met Ser Ile Leu Lys Gln
                    215                           220                           225
Cys Gly Gln Glu Ala Ala Leu Leu Ile Glu Tyr Asp Val Ser Val
                    230                           235                           240
Met Asp Ser Val Ala Thr Ala Ser Gly Pro Leu Leu Val Glu Val
                    245                           250                           255
Ala Lys Thr Pro Gly Ala Ser Leu Gly Val Ala Leu Thr Thr Ser
                    260                           265                           270
Met Cys Cys Asn Lys Gln Val Ile Val Ile Asp Lys Ile Lys Ser
                    275                           280                           285
Ala Ser Ile Ala Asp Arg Cys Gly Ala Leu His Val Gly Asp His
                    290                           295                           300
Ile Leu Ser Ile Asp Gly Thr Ser Met Glu Tyr Cys Thr Leu Ala
                    305                           310                           315
Glu Ala Thr Gln Phe Leu Ala Asn Thr Thr Asp Gln Val Lys Leu
                    320                           325                           330
Glu Ile Leu Pro His His Gln Thr Arg Leu Ala Leu Lys Gly Pro
                    335                           340                           345
Asp His Val Lys Ile Gln Arg Ser Asp Arg Gln Leu Thr Trp Asp
                    350                           355                           360
Ser Trp Ala Ser Asn His Ser Ser Leu His Thr Asn His His Tyr
                    365                           370                           375
Asn Thr Tyr His Pro Asp His Cys Arg Val Pro Ala Leu Thr Phe
                    380                           385                           390
Pro Lys Ala Pro Pro Pro Asn Ser Pro Pro Ala Leu Val Ser Ser
                    395                           400                           405
Ser Phe Ser Pro Thr Ser Met Ser Ala Tyr Ser Leu Ser Ser Leu
                    410                           415                           420
Asn Met Gly Thr Leu Pro Arg Ser Leu Tyr Ser Thr Ser Pro Arg
                    425                           430                           435
Gly Thr Met Met Arg Arg Arg Leu Lys Lys Lys Asp Phe Lys Ser
                    440                           445                           450
Ser Leu Ser Leu Ala Ser Ser Thr Val Gly Leu Ala Gly Gln Val
                    455                           460                           465
Val His Thr Glu Thr Thr Glu Val Val Leu Thr Ala Asp Pro Val
                    470                           475                           480
Thr Gly Phe Gly Ile Gln Leu Gln Gly Ser Val Phe Ala Thr Glu
                    485                           490                           495
Thr Leu Ser Ser Pro Pro Leu Ile Ser Tyr Ile Glu Ala Asp Ser
                    500                           505                           510
Pro Ala Glu Arg Cys Gly Val Leu Gln Ile Gly Asp Arg Val Met
                    515                           520                           525
Ala Ile Asn Gly Ile Pro Thr Glu Asp Ser Thr Phe Glu Glu Ala
                    530                           535                           540
Ser Gln Leu Leu Arg Asp Ser Ser Ile Thr Ser Lys Val Thr Leu
                    545                           550                           555
Glu Ile Glu Phe Asp Val Ala Glu Ser Val Ile Pro Ser Ser Gly
                    560                           565                           570
Thr Phe His Val Lys Leu Pro Lys Lys His Asn Val Glu Leu Gly
                    575                           580                           585
Ile Thr Ile Ser Ser Pro Ser Ser Arg Lys Pro Gly Asp Pro Leu
                    590                           595                           600
Val Ile Ser Asp Ile Lys Lys Gly Ser Val Ala His Arg Thr Gly
                    605                           610                           615
Thr Leu Glu Leu Gly Asp Lys Leu Leu Ala Ile Asp Asn Ile Arg
                    620                           625                           630
Leu Asp Asn Cys Ser Met Glu Asp Ala Val Gln Ile Leu Gln Gln
                    635                           640                           645
```

Cys Glu Asp Leu Val Lys Leu Lys Ile Arg Lys Asp Glu Asp Asn
650 655 660
Ser Asp Glu Gln Glu Ser Ser Gly Ala Ile Ile Tyr Thr Val Glu
665 670 675
Leu Lys Arg Tyr Gly Gly Pro Leu Gly Ile Thr Ile Ser Gly Thr
680 685 690
Glu Glu Pro Phe Asp Pro Ile Ile Ile Ser Ser Leu Thr Lys Gly
695 700 705
Gly Leu Ala Glu Arg Thr Gly Ala Ile His Ile Gly Asp Arg Ile
710 715 720
Leu Ala Ile Asn Ser Ser Ser Leu Lys Gly Lys Pro Leu Ser Glu
725 730 735
Ala Ile His Leu Leu Gln Met Ala Gly Glu Thr Val Thr Leu Lys
740 745 750
Ile Lys Lys Gln Thr Asp Ala Gln Ser Ala Ser Ser Pro Lys Lys
755 760 765
Phe Pro Ile Ser Ser His Leu Ser Asp Leu Gly Asp Val Glu Glu
770 775 780
Asp Ser Ser Pro Ala Gln Lys Pro Gly Lys Leu Ser Asp Met Tyr
785 790 795
Pro Ser Thr Val Pro Ser Val Asp Ser Ala Val Asp Ser Trp Asp
800 805 810
Gly Ser Ala Ile Asp Thr Ser Tyr Gly Thr Glu Gly Thr Ser Phe
815 820 825
Gln Ala Ser Gly Tyr Asn Phe Asn Thr Tyr Asp Trp Arg Ser Pro
830 835 840
Lys Gln Arg Gly Ser Leu Ser Pro Val Thr Lys Pro Arg Ser Gln
845 850 855
Thr Tyr Pro Asp Val Gly Leu Ser Tyr Glu Asp Trp Asp Arg Ser
860 865 870
Thr Ala Ser Gly Phe Ala Gly Ala Ala Asp Ser Ala Glu Thr Glu
875 880 885
Gln Glu Glu Asn Phe Trp Ser Gln Ala Leu Glu Asp Leu Glu Thr
890 895 900
Cys Gly Gln Ser Gly Ile Leu Arg Glu Leu Glu Ala Thr Ile Met
905 910 915
Ser Gly Ser Thr Met Ser Leu Asn His Glu Ala Pro Thr Pro Arg
920 925 930
Ser Gln Leu Gly Arg Gln Ala Ser Phe Gln Glu Arg Ser Ser Ser
935 940 945
Arg Pro His Tyr Ser Gln Thr Thr Arg Ser Asn Thr Leu Pro Ser
950 955 960
Asp Val Gly Arg Lys Ser Val Thr Leu Arg Lys Met Lys Gln Glu
965 970 975
Ile Lys Glu Ile Met Ser Pro Thr Pro Val Glu Leu His Lys Val
980 985 990
Thr Leu Tyr Lys Asp Ser Asp Met Glu Asp Phe Gly Phe Ser Val
995 1000 1005
Ala Asp Gly Leu Leu Glu Lys Gly Val Tyr Val Lys Asn Ile Arg
1010 1015 1020
Pro Ala Gly Pro Gly Asp Leu Gly Gly Leu Lys Pro Tyr Asp Arg
1025 1030 1035
Leu Leu Gln Val Asn His Val Arg Thr Arg Asp Phe Asp Cys Cys
1040 1045 1050
Leu Val Val Pro Leu Ile Ala Glu Ser Gly Asn Lys Leu Asp Leu
1055 1060 1065
Val Ile Ser Arg Asn Pro Leu Ala Ser Gln Lys Ser Ile Asp Gln
1070 1075 1080
Gln Ser Leu Pro Gly Asp Trp Ser Glu Gln Asn Ser Ala Phe Phe
1085 1090 1095
Gln Gln Pro Ser His Gly Gly Asn Leu Glu Thr Arg Glu Pro Thr
1100 1105 1110
Asn Thr Leu

<210> 37
<211> 511
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1338201CD1

<400> 37
Met Ser Arg Gly Pro Glu Glu Val Asn Arg Leu Thr Glu Ser Thr
1 5 10 15
Tyr Arg Asn Val Met Glu Gln Phe Asn Pro Gly Leu Arg Asn Leu
20 25 30
Ile Asn Leu Gly Lys Asn Tyr Glu Lys Ala Val Asn Ala Met Ile
35 40 45
Leu Ala Gly Lys Ala Tyr Tyr Asp Gly Val Ala Lys Ile Gly Glu
50 55 60
Ile Ala Thr Gly Ser Pro Val Ser Thr Glu Leu Gly His Val Leu
65 70 75
Ile Glu Ile Ser Ser Thr His Lys Lys Leu Asn Glu Ser Leu Asp
80 85 90
Glu Asn Phe Lys Lys Phe His Lys Glu Ile Ile His Glu Leu Glu
95 100 105
Lys Lys Ile Glu Leu Asp Val Lys Tyr Met Asn Ala Thr Leu Lys
110 115 120
Arg Tyr Gln Thr Glu His Lys Asn Lys Leu Glu Ser Leu Glu Lys
125 130 135
Ser Gln Ala Glu Leu Lys Lys Ile Arg Arg Lys Ser Gln Gly Ser
140 145 150
Arg Asn Ala Leu Lys Tyr Glu His Lys Glu Ile Glu Tyr Val Glu
155 160 165
Thr Val Thr Ser Arg Gln Ser Glu Ile Gln Lys Phe Ile Ala Asp
170 175 180
Gly Cys Lys Glu Ala Leu Leu Glu Glu Lys Arg Arg Phe Cys Phe
185 190 195
Leu Val Asp Lys His Cys Gly Phe Ala Asn His Ile His Tyr Tyr
200 205 210
His Leu Gln Ser Ala Glu Leu Leu Asn Ser Lys Leu Pro Arg Trp
215 220 225
Gln Glu Thr Cys Val Asp Ala Ile Lys Val Pro Glu Lys Ile Met
230 235 240
Asn Met Ile Glu Glu Ile Lys Thr Pro Ala Ser Thr Pro Val Ser
245 250 255
Gly Thr Pro Gln Ala Ser Pro Met Ile Glu Arg Ser Asn Val Val
260 265 270
Arg Lys Asp Tyr Asp Thr Leu Ser Lys Cys Ser Pro Lys Met Pro
275 280 285
Pro Ala Pro Ser Gly Arg Ala Tyr Thr Ser Pro Leu Ile Asp Met
290 295 300
Phe Asn Asn Pro Ala Thr Ala Ala Pro Asn Ser Gln Arg Val Asn
305 310 315
Asn Ser Thr Gly Thr Ser Glu Asp Pro Ser Leu Gln Arg Ser Val
320 325 330
Ser Val Ala Thr Gly Leu Asn Met Met Lys Lys Gln Lys Val Lys
335 340 345
Thr Ile Phe Pro His Thr Ala Gly Ser Asn Lys Thr Leu Leu Ser
350 355 360
Phe Ala Gln Gly Asp Val Ile Thr Leu Leu Ile Pro Glu Glu Lys
365 370 375
Asp Gly Trp Leu Tyr Gly Glu His Asp Val Ser Lys Ala Arg Gly
380 385 390
Trp Phe Pro Ser Ser Tyr Thr Lys Leu Leu Glu Glu Asn Glu Thr
395 400 405
Glu Ala Val Thr Val Pro Thr Pro Ser Pro Thr Pro Val Arg Ser
410 415 420
Ile Ser Thr Val Asn Leu Ser Glu Asn Ser Ser Val Val Ile Pro
425 430 435
Pro Pro Asp Tyr Leu Glu Cys Leu Ser Met Gly Ala Ala Ala Asp
440 445 450

Arg Arg Ala Asp Ser Ala Arg Thr Thr Ser Thr Phe Lys Ala Pro
455 460 465
Ala Ser Lys Pro Glu Thr Ala Ala Pro Asn Asp Ala Asn Gly Thr
470 475 480
Ala Lys Pro Pro Phe Leu Ser Gly Glu Asn Pro Phe Ala Thr Val
485 490 495
Lys Leu Arg Pro Thr Val Thr Asn Asp Arg Ser Ala Pro Ile Ile
500 505 510
Arg

<210> 38
<211> 1177
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1405141CD1

<400> 38
Met Thr Thr Ile Leu Lys Pro Ser Ala Asp Phe Leu Thr Ser Asn
1 5 10 15
Lys Leu Leu Lys Tyr Ser Trp Phe Phe Phe Asp Val Leu Ile Lys
20 25 30
Ser Met Ala Gln His Leu Ile Glu Asn Ser Lys Val Lys Leu Leu
35 40 45
Arg Asn Gln Arg Phe Pro Ala Ser Tyr His His Ala Val Glu Thr
50 55 60
Val Val Asn Met Leu Met Pro His Ile Thr Gln Lys Phe Arg Asp
65 70 75
Asn Pro Glu Ala Ser Lys Asn Ala Asn His Ser Leu Ala Val Phe
80 85 90
Ile Lys Arg Cys Phe Thr Phe Met Asp Arg Gly Phe Val Phe Lys
95 100 105
Gln Ile Asn Asn Tyr Ile Ser Cys Phe Ala Pro Gly Asp Pro Lys
110 115 120
Thr Leu Phe Glu Tyr Lys Phe Glu Phe Leu Arg Val Val Cys Asn
125 130 135
His Glu His Tyr Ile Pro Leu Asn Leu Pro Met Pro Phe Gly Lys
140 145 150
Gly Arg Ile Gln Arg Tyr Gln Asp Leu Gln Leu Asp Tyr Ser Leu
155 160 165
Thr Asp Glu Phe Cys Arg Asn His Phe Leu Val Gly Leu Leu Leu
170 175 180
Arg Glu Val Gly Thr Ala Leu Gln Glu Phe Arg Glu Val Arg Leu
185 190 195
Ile Ala Ile Ser Val Leu Lys Asn Leu Leu Ile Lys His Ser Phe
200 205 210
Asp Asp Arg Tyr Ala Ser Arg Ser His Gln Ala Arg Ile Ala Thr
215 220 225
Leu Tyr Leu Pro Leu Phe Gly Leu Leu Ile Glu Asn Val Gln Arg
230 235 240
Ile Asn Val Arg Asp Val Ser Pro Phe Pro Val Asn Ala Gly Met
245 250 255
Thr Val Lys Asp Glu Ser Leu Ala Leu Pro Ala Val Asn Pro Leu
260 265 270
Val Thr Pro Gln Lys Gly Ser Thr Leu Asp Asn Ser Leu His Lys
275 280 285
Asp Leu Leu Gly Ala Ile Ser Gly Ile Ala Ser Pro Tyr Thr Thr
290 295 300
Ser Thr Pro Asn Ile Asn Ser Val Arg Asn Ala Asp Ser Arg Gly
305 310 315
Ser Leu Ile Ser Thr Asp Ser Gly Asn Ser Leu Pro Glu Arg Asn
320 325 330
Ser Glu Lys Ser Asn Ser Leu Asp Lys His Gln Gln Ser Ser Thr
335 340 345
Leu Gly Asn Ser Val Val Arg Cys Asp Lys Leu Asp Gln Ser Glu

350 355 360
Ile Lys Ser Leu Leu Met Cys Phe Leu Tyr Ile Leu Lys Ser Met
365 370 375
Ser Asp Asp Ala Leu Phe Thr Tyr Trp Asn Lys Ala Ser Thr Ser
380 385 390
Glu Leu Met Asp Phe Phe Thr Ile Ser Glu Val Cys Leu His Gln
395 400 405
Phe Gln Tyr Met Gly Lys Arg Tyr Ile Ala Ser Val Arg Lys Ile
410 415 420
Ser Ser Val Leu Gly Ile Ser Val Asp Asn Gly Tyr Gly His Ser
425 430 435
Asp Ala Asp Val Leu His Gln Ser Leu Leu Glu Ala Asn Ile Ala
440 445 450
Thr Glu Val Cys Leu Thr Ala Leu Asp Thr Leu Ser Leu Phe Thr
455 460 465
Leu Ala Phe Lys Asn Gln Leu Leu Ala Asp His Gly His Asn Pro
470 475 480
Leu Met Lys Lys Val Phe Asp Val Tyr Leu Cys Phe Leu Gln Lys
485 490 495
His Gln Ser Glu Thr Ala Leu Lys Asn Val Phe Thr Ala Leu Arg
500 505 510
Ser Leu Ile Tyr Lys Phe Pro Ser Thr Phe Tyr Glu Gly Arg Ala
515 520 525
Asp Met Cys Ala Ala Leu Cys Tyr Glu Ile Leu Lys Cys Cys Asn
530 535 540
Ser Lys Leu Ser Ser Ile Arg Thr Glu Ala Ser Gln Leu Leu Tyr
545 550 555
Phe Leu Met Arg Asn Asn Phe Asp Tyr Thr Gly Lys Lys Ser Phe
560 565 570
Val Arg Thr His Leu Gln Val Ile Ile Ser Val Ser Gln Leu Ile
575 580 585
Ala Asp Val Val Gly Ile Gly Gly Thr Arg Phe Gln Gln Ser Leu
590 595 600
Ser Ile Ile Asn Asn Cys Ala Asn Ser Asp Arg Leu Ile Lys His
605 610 615
Thr Ser Phe Ser Ser Asp Val Lys Asp Leu Thr Lys Arg Ile Arg
620 625 630
Thr Val Leu Met Ala Thr Ala Gln Met Lys Glu His Glu Asn Asp
635 640 645
Pro Glu Met Leu Val Asp Leu Gln Tyr Ser Leu Ala Lys Ser Tyr
650 655 660
Ala Ser Thr Pro Glu Leu Arg Lys Thr Trp Leu Asp Ser Met Ala
665 670 675
Arg Ile His Val Lys Asn Gly Asp Leu Ser Glu Ala Ala Met Cys
680 685 690
Tyr Val His Val Thr Ala Leu Val Ala Glu Tyr Leu Thr Arg Lys
695 700 705
Gly Val Phe Arg Gln Gly Cys Thr Ala Phe Arg Val Ile Thr Pro
710 715 720
Asn Ile Asp Glu Glu Ala Ser Met Met Glu Asp Val Gly Met Gln
725 730 735
Asp Val His Phe Asn Glu Asp Val Leu Met Glu Leu Leu Glu Gln
740 745 750
Cys Ala Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Leu Ile Ala
755 760 765
Asp Ile Tyr Lys Leu Ile Ile Pro Ile Tyr Glu Lys Arg Arg Asp
770 775 780
Phe Glu Arg Leu Ala His Leu Tyr Asp Thr Leu His Arg Ala Tyr
785 790 795
Ser Lys Val Thr Glu Val Met His Ser Gly Arg Ser Val Leu Gly
800 805 810
Thr Tyr Phe Arg Val Ala Phe Phe Gly Gln Gly Phe Phe Glu Asp
815 820 825
Glu Asp Gly Lys Glu Tyr Ile Tyr Lys Glu Pro Lys Leu Thr Pro
830 835 840
Leu Ser Glu Ile Ser Gln Arg Leu Leu Lys Leu Tyr Ser Asp Lys
845 850 855

Phe Gly Ser Glu Asn Val Lys Met Ile Gln Asp Ser Gly Lys Val
860 865 870
Asn Pro Lys Asp Leu Asp Ser Lys Tyr Ala Tyr Ile Gln Val Thr
875 880 885
His Val Ile Pro Phe Phe Asp Glu Lys Glu Leu Gln Glu Arg Lys
890 895 900
Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg Phe Met Phe Glu
905 910 915
Met Pro Phe Thr Gln Thr Gly Lys Arg Gln Gly Gly Val Glu Glu
920 925 930
Gln Cys Lys Arg Arg Thr Ile Leu Thr Ala Ile His Cys Phe Pro
935 940 945
Tyr Val Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr Asp
950 955 960
Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val
965 970 975
Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile
980 985 990
Lys Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn
995 1000 1005
Ala Gly Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn
1010 1015 1020
Thr Lys Arg Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val
1025 1030 1035
Phe Arg Gln Phe Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn
1040 1045 1050
Glu Arg Leu Ile Lys Glu Asp Gln Leu Glu Tyr Gln Glu Glu Met
1055 1060 1065
Lys Ala Asn Tyr Arg Glu Met Ala Lys Glu Leu Ser Glu Ile Met
1070 1075 1080
His Glu Gln Ile Cys Pro Leu Glu Asp Glu Asp Glu Arg Leu Thr
1085 1090 1095
Glu Phe Pro Ser His Leu Gln Arg His Gln Trp Asp Ser Asn Lys
1100 1105 1110
His Asn Gly Ser Arg Asp Asp Gln Leu Val Phe Gly Arg Val Ile
1115 1120 1125
Thr Ser His Gly Pro Cys Val Gly Thr Cys Phe Val Ile Cys Lys
1130 1135 1140
Leu Arg Met Leu Ser Lys Ala Asn His Trp Gly Asp Arg Ala Gln
1145 1150 1155
Gly Gly Pro Arg Gly Arg Gly Glu Lys Gly Asn Lys Glu Gln Arg
1160 1165 1170
Tyr Phe Leu Thr Asp Phe Leu
1175

<210> 39
<211> 665
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1686305CD1

<400> 39
Met Thr Ser Ala Asn Lys Ala Ile Glu Leu Gln Leu Gln Val Lys
1 5 10 15
Gln Asn Ala Glu Glu Leu Gln Asp Phe Met Arg Asp Leu Glu Asn
20 25 30
Trp Glu Lys Asp Ile Lys Gln Lys Asp Met Glu Leu Arg Arg Gln
35 40 45
Asn Gly Val Pro Glu Glu Asn Leu Pro Pro Ile Arg Asn Gly Asn
50 55 60
Phe Arg Lys Lys Lys Lys Gly Lys Ala Lys Glu Ser Ser Lys Lys
65 70 75
Thr Arg Glu Glu Asn Thr Lys Asn Arg Ile Lys Ser Tyr Asp Tyr
80 85 90
Glu Ala Trp Ala Lys Leu Asp Val Asp Arg Ile Leu Asp Glu Leu

```
                    95                     100                    105
Asp Lys Asp Asp Ser Thr His Glu Ser Leu Ser Gln Glu Ser Glu
                110                    115                    120
Ser Glu Glu Asp Gly Ile His Val Asp Ser Gln Lys Ala Leu Val
                125                    130                    135
Leu Lys Glu Lys Gly Asn Lys Tyr Phe Lys Gln Gly Lys Tyr Asp
                140                    145                    150
Glu Ala Ile Asp Cys Tyr Thr Lys Gly Met Asp Ala Asp Pro Tyr
                155                    160                    165
Asn Pro Val Leu Pro Thr Asn Arg Ala Ser Ala Tyr Phe Arg Leu
                170                    175                    180
Lys Lys Phe Ala Val Ala Glu Ser Asp Cys Asn Leu Ala Val Ala
                185                    190                    195
Leu Asn Arg Ser Tyr Thr Lys Ala Tyr Ser Arg Arg Gly Ala Ala
                200                    205                    210
Arg Phe Ala Leu Gln Lys Leu Glu Glu Ala Lys Lys Asp Tyr Glu
                215                    220                    225
Arg Val Leu Glu Leu Glu Pro Asn Asn Phe Glu Ala Thr Asn Glu
                230                    235                    240
Leu Arg Lys Ile Ser Gln Ala Leu Ala Ser Lys Glu Asn Ser Tyr
                245                    250                    255
Pro Lys Glu Ala Asp Ile Val Ile Lys Ser Thr Glu Gly Glu Arg
                260                    265                    270
Lys Gln Ile Glu Ala Gln Gln Asn Lys Gln Gln Ala Ile Ser Glu
                275                    280                    285
Lys Asp Arg Gly Asn Gly Phe Phe Lys Glu Gly Lys Tyr Glu Arg
                290                    295                    300
Ala Ile Glu Cys Tyr Thr Arg Gly Ile Ala Ala Asp Gly Ala Asn
                305                    310                    315
Ala Leu Leu Pro Ala Asn Arg Ala Met Ala Tyr Leu Lys Ile Gln
                320                    325                    330
Lys Tyr Glu Glu Ala Glu Lys Asp Cys Thr Gln Ala Ile Leu Leu
                335                    340                    345
Asp Gly Ser Tyr Ser Lys Ala Phe Ala Arg Arg Gly Thr Ala Arg
                350                    355                    360
Thr Phe Leu Gly Lys Leu Asn Glu Ala Lys Gln Asp Phe Glu Thr
                365                    370                    375
Val Leu Leu Leu Glu Pro Gly Asn Lys Gln Ala Val Thr Glu Leu
                380                    385                    390
Ser Lys Ile Lys Lys Glu Leu Ile Glu Lys Gly His Trp Asp Asp
                395                    400                    405
Val Phe Leu Asp Ser Thr Gln Arg Gln Asn Val Val Lys Pro Ile
                410                    415                    420
Asp Asn Pro Pro His Pro Gly Ser Thr Lys Pro Leu Lys Lys Val
                425                    430                    435
Ile Ile Glu Glu Thr Gly Asn Leu Ile Gln Thr Ile Asp Val Pro
                440                    445                    450
Asp Ser Thr Thr Ala Ala Ala Pro Glu Asn Asn Pro Ile Asn Leu
                455                    460                    465
Ala Asn Val Ile Ala Ala Thr Gly Thr Thr Ser Lys Lys Asn Ser
                470                    475                    480
Ser Gln Asp Asp Leu Phe Pro Thr Ser Asp Thr Pro Arg Ala Lys
                485                    490                    495
Val Leu Lys Ile Glu Glu Val Ser Asp Thr Ser Ser Leu Gln Pro
                500                    505                    510
Gln Ala Ser Leu Lys Gln Asp Val Cys Gln Ser Tyr Ser Glu Lys
                515                    520                    525
Met Pro Ile Glu Ile Glu Gln Lys Pro Ala Gln Phe Ala Thr Thr
                530                    535                    540
Val Leu Pro Pro Ile Pro Ala Asn Ser Phe Gln Leu Glu Ser Asp
                545                    550                    555
Phe Arg Gln Leu Lys Ser Ser Pro Asp Met Leu Tyr Gln Tyr Leu
                560                    565                    570
Lys Gln Ile Glu Pro Ser Leu Tyr Pro Lys Leu Phe Gln Lys Asn
                575                    580                    585
Leu Asp Pro Asp Val Phe Asn Gln Ile Val Lys Ile Leu His Asp
                590                    595                    600
```

Phe Tyr Ile Glu Lys Glu Lys Pro Leu Leu Ile Phe Glu Ile Leu
605 610 615
Gln Arg Leu Ser Glu Leu Lys Arg Phe Asp Met Ala Val Met Phe
620 625 630
Met Ser Glu Thr Glu Lys Lys Ile Ala Arg Ala Leu Phe Asn His
635 640 645
Ile Asp Lys Ser Gly Leu Lys Asp Ser Ser Val Glu Glu Leu Lys
650 655 660
Lys Arg Tyr Gly Gly
665
<210> 40
<211> 125
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1688972CD1

<400> 40
Met Leu Asp Leu Gln Lys Gln Leu Gly Arg Phe Pro Gly Thr Phe
1 5 10 15
Val Gly Thr Thr Glu Pro Ala Ser Pro Pro Leu Ser Ser Thr Ser
20 25 30
Pro Thr Thr Ala Ala Ala Thr Met Pro Val Val Pro Ser Val Ala
35 40 45
Ser Leu Ala Pro Pro Gly Glu Ala Ser Leu Cys Leu Glu Glu Val
50 55 60
Ala Pro Pro Ala Ser Gly Thr Arg Lys Ala Arg Val Leu Tyr Asp
65 70 75
Tyr Glu Ala Ala Asp Ser Ser Glu Leu Ala Leu Leu Ala Asp Glu
80 85 90
Leu Ile Thr Val Tyr Ser Leu Pro Gly Met Asp Pro Asp Trp Leu
95 100 105
Ile Gly Glu Arg Gly Asn Lys Lys Gly Lys Val Pro Val Thr Tyr
110 115 120
Leu Glu Leu Leu Ser
125
<210> 41
<211> 366
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812494CD1

<400> 41
Met Cys Tyr Phe Tyr Leu Gly Asp Lys Ile Lys Thr Ile Ser Phe
1 5 10 15
Gln Ala Phe Ile Leu Met His Leu Leu Leu Pro Ser Glu Tyr Ser
20 25 30
Leu Asp Gly Phe His Met Ser Gly Phe Ser Leu Gly Ser Gly Ser
35 40 45
Glu Gly Glu Asp Gly Phe Gln Val Glu Leu Glu Leu Val Glu Leu
50 55 60
Thr Val Gly Thr Leu Asp Leu Cys Glu Ser Glu Val Leu Pro Lys
65 70 75
Arg Arg Arg Arg Lys Arg Asn Lys Lys Glu Lys Ser Arg Asp Gln
80 85 90
Glu Ala Gly Ala His Arg Thr Leu Leu Gln Gln Thr Gln Glu Glu
95 100 105
Glu Pro Ser Thr Gln Ser Ser Gln Ala Val Ala Ala Pro Leu Gly
110 115 120
Pro Leu Leu Asp Glu Ala Lys Ala Pro Gly Gln Pro Glu Leu Trp
125 130 135
Asn Ala Leu Leu Ala Ala Cys Arg Ala Gly Asp Val Gly Val Leu

```
                    140                           145                           150
Lys Leu Gln Leu Ala Pro Ser Pro Ala Asp Pro Arg Val Leu Ser
                    155                           160                           165
Leu Leu Ser Ala Pro Leu Gly Ser Gly Gly Phe Thr Leu Leu His
                    170                           175                           180
Ala Ala Ala Ala Ala Gly Arg Gly Ser Val Val Arg Leu Leu Leu
                    185                           190                           195
Glu Ala Gly Ala Asp Pro Thr Val Gln Asp Ser Arg Ala Arg Pro
                    200                           205                           210
Pro Tyr Thr Val Ala Ala Asp Lys Ser Thr Arg Asn Glu Phe Arg
                    215                           220                           225
Arg Phe Met Glu Lys Asn Pro Asp Ala Tyr Asp Tyr Asn Lys Ala
                    230                           235                           240
Gln Val Pro Gly Pro Leu Thr Pro Glu Met Glu Ala Arg Gln Ala
                    245                           250                           255
Thr Arg Lys Arg Glu Gln Lys Ala Ala Arg Arg Gln Arg Glu Glu
                    260                           265                           270
Gln Gln Gln Arg Gln Gln Glu Gln Glu Glu Arg Glu Arg Glu Glu
                    275                           280                           285
Gln Arg Arg Phe Ala Ala Leu Ser Asp Arg Glu Lys Arg Ala Leu
                    290                           295                           300
Ala Ala Glu Arg Arg Leu Ala Ala Gln Leu Gly Ala Pro Thr Ser
                    305                           310                           315
Pro Ile Pro Asp Ser Ala Ile Val Asn Thr Arg Arg Cys Trp Ser
                    320                           325                           330
Cys Gly Ala Ser Leu Gln Gly Leu Thr Pro Phe His Tyr Leu Asp
                    335                           340                           345
Phe Ser Phe Cys Ser Thr Arg Cys Leu Gln Asp His Arg Arg Gln
                    350                           355                           360
Ala Gly Arg Pro Ser Ser
                    365
<210> 42
<211> 173
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2013853CD1

<400> 42
Met Ser Thr Met Gly Asn Glu Ala Ser Tyr Pro Ala Glu Met Cys
  1                 5                            10                            15
Ser His Phe Asp Asn Asp Glu Ile Lys Arg Leu Gly Arg Arg Phe
                    20                            25                            30
Lys Lys Leu Asp Leu Asp Lys Ser Gly Ser Leu Ser Val Glu Glu
                    35                            40                            45
Phe Met Ser Leu Pro Glu Leu Arg His Asn Pro Leu Val Arg Arg
                    50                            55                            60
Val Ile Asp Val Phe Asp Thr Asp Gly Asp Gly Glu Val Asp Phe
                    65                            70                            75
Lys Glu Phe Ile Leu Gly Thr Ser Gln Phe Ser Val Lys Gly Asp
                    80                            85                            90
Glu Glu Gln Lys Leu Arg Phe Ala Phe Ser Ile Tyr Asp Met Asp
                    95                           100                           105
Lys Asp Gly Tyr Ile Ser Asn Gly Glu Leu Phe Gln Val Leu Lys
                    110                           115                           120
Met Met Val Gly Asn Asn Leu Thr Asp Trp Gln Leu Gln Gln Leu
                    125                           130                           135
Val Asp Lys Thr Ile Ile Ile Leu Asp Lys Asp Gly Asp Gly Lys
                    140                           145                           150
Ile Ser Phe Glu Glu Phe Ser Ala Val Val Arg Asp Leu Glu Ile
                    155                           160                           165
His Lys Lys Leu Val Leu Ile Val
                    170
<210> 43
<211> 761
```

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2284925CD1

<400> 43
Met Arg Leu Thr Gln Asp Pro Ile Gln Val Leu Leu Ile Phe Ala
1 5 10 15
Lys Glu Asp Ser Gln Ser Asp Gly Phe Trp Trp Ala Cys Asp Arg
20 25 30
Ala Gly Tyr Arg Cys Asn Ile Ala Arg Thr Pro Glu Ser Ala Leu
35 40 45
Glu Cys Phe Leu Asp Lys His His Glu Ile Ile Val Ile Asp His
50 55 60
Arg Gln Thr Gln Asn Phe Asp Ala Glu Ala Val Cys Arg Ser Ile
65 70 75
Arg Ala Thr Asn Pro Ser Glu His Thr Val Ile Leu Ala Val Val
80 85 90
Ser Arg Val Ser Asp Asp His Glu Glu Ala Ser Val Leu Pro Leu
95 100 105
Leu His Ala Gly Phe Asn Arg Arg Phe Met Glu Asn Ser Ser Ile
110 115 120
Ile Ala Cys Tyr Asn Glu Leu Ile Gln Ile Glu His Gly Glu Val
125 130 135
Arg Ser Gln Phe Lys Leu Arg Ala Cys Asn Ser Val Phe Thr Ala
140 145 150
Leu Asp His Cys His Glu Ala Ile Glu Ile Thr Ser Asp Asp His
155 160 165
Val Ile Gln Tyr Val Asn Pro Ala Phe Glu Arg Met Met Gly Tyr
170 175 180
His Lys Gly Glu Leu Leu Gly Lys Glu Leu Ala Asp Leu Pro Lys
185 190 195
Ser Asp Lys Asn Arg Ala Asp Leu Leu Asp Thr Ile Asn Thr Cys
200 205 210
Ile Lys Lys Gly Lys Glu Trp Gln Gly Val Tyr Tyr Ala Arg Arg
215 220 225
Lys Ser Gly Asp Ser Ile Gln Gln His Val Lys Ile Thr Pro Val
230 235 240
Ile Gly Gln Gly Gly Lys Ile Arg His Phe Val Ser Leu Lys Lys
245 250 255
Leu Cys Cys Thr Thr Asp Asn Asn Lys Gln Ile His Lys Ile His
260 265 270
Arg Asp Ser Gly Asp Asn Ser Gln Thr Glu Pro His Ser Phe Arg
275 280 285
Tyr Lys Asn Arg Arg Lys Glu Ser Ile Asp Val Lys Ser Ile Ser
290 295 300
Ser Arg Gly Ser Asp Ala Pro Ser Leu Gln Asn Arg Arg Tyr Pro
305 310 315
Ser Met Ala Arg Ile His Ser Met Thr Ile Glu Ala Pro Ile Thr
320 325 330
Lys Val Ile Asn Ile Ile Asn Ala Ala Gln Glu Asn Ser Pro Val
335 340 345
Thr Val Ala Glu Ala Leu Asp Arg Val Leu Glu Ile Leu Arg Thr
350 355 360
Thr Glu Leu Tyr Ser Pro Gln Leu Gly Thr Lys Asp Glu Asp Pro
365 370 375
His Thr Ser Asp Leu Val Gly Gly Leu Met Thr Asp Gly Leu Arg
380 385 390
Arg Leu Ser Gly Asn Glu Tyr Val Phe Thr Lys Asn Val His Gln
395 400 405
Ser His Ser His Leu Ala Met Pro Ile Thr Ile Asn Asp Val Pro
410 415 420
Pro Cys Ile Ser Gln Leu Leu Asp Asn Glu Glu Ser Trp Asp Phe
425 430 435
Asn Ile Phe Glu Leu Glu Ala Ile Thr His Lys Arg Pro Leu Val

```
                    440                                     445                                     450
Tyr Leu Gly Leu Lys Val Phe Ser Arg Phe Gly Val Cys Glu Phe
                    455                 460                 465
Leu Asn Cys Ser Glu Thr Thr Leu Arg Ala Trp Phe Gln Val Ile
                    470                 475                 480
Glu Ala Asn Tyr His Ser Ser Asn Ala Tyr His Asn Ser Thr His
                    485                 490                 495
Ala Ala Asp Val Leu His Ala Thr Ala Phe Phe Leu Gly Lys Glu
                    500                 505                 510
Arg Val Lys Gly Ser Leu Asp Gln Leu Asp Glu Val Ala Ala Leu
                    515                 520                 525
Ile Ala Ala Thr Val His Asp Val Asp His Pro Gly Arg Thr Asn
                    530                 535                 540
Ser Phe Leu Cys Asn Ala Gly Ser Glu Leu Ala Val Leu Tyr Asn
                    545                 550                 555
Asp Thr Ala Val Leu Glu Ser His His Thr Ala Leu Ala Phe Gln
                    560                 565                 570
Leu Thr Val Lys Asp Thr Lys Cys Asn Ile Phe Lys Asn Ile Asp
                    575                 580                 585
Arg Asn His Tyr Arg Thr Leu Arg Gln Ala Ile Ile Asp Met Val
                    590                 595                 600
Leu Ala Thr Glu Met Thr Lys His Phe Glu His Val Asn Lys Phe
                    605                 610                 615
Val Asn Ser Ile Asn Lys Pro Met Ala Ala Glu Ile Glu Gly Ser
                    620                 625                 630
Asp Cys Glu Cys Asn Pro Ala Gly Lys Asn Phe Pro Glu Asn Gln
                    635                 640                 645
Ile Leu Ile Lys Arg Met Met Ile Lys Cys Ala Asp Val Ala Asn
                    650                 655                 660
Pro Cys Arg Pro Leu Asp Leu Cys Ile Glu Trp Ala Gly Arg Ile
                    665                 670                 675
Ser Glu Glu Tyr Phe Ala Gln Thr Asp Glu Glu Lys Arg Gln Gly
                    680                 685                 690
Leu Pro Val Val Met Pro Val Phe Asp Arg Asn Thr Cys Ser Ile
                    695                 700                 705
Pro Lys Ser Gln Ile Ser Phe Ile Asp Tyr Phe Ile Thr Asp Met
                    710                 715                 720
Phe Asp Ala Trp Asp Ala Phe Ala His Leu Pro Ala Leu Met Gln
                    725                 730                 735
His Leu Ala Asp Asn Tyr Lys His Trp Lys Thr Leu Asp Asp Leu
                    740                 745                 750
Lys Cys Lys Ser Leu Arg Leu Pro Ser Asp Ser
                    755                 760
```

<210> 44
<211> 249
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2376728CD1

<400> 44

```
Met Val Asp Arg Leu Ala Asn Ser Glu Ala Asn Thr Arg Arg Ile
 1               5                  10                  15
Ser Ile Val Glu Asn Cys Phe Gly Ala Ala Gly Gln Pro Leu Thr
                20                  25                  30
Ile Pro Gly Arg Val Leu Ile Gly Glu Gly Val Leu Thr Lys Leu
                35                  40                  45
Cys Arg Lys Lys Pro Lys Ala Arg Gln Phe Phe Leu Phe Asn Asp
                50                  55                  60
Ile Leu Val Tyr Gly Asn Ile Val Ile Gln Lys Lys Lys Tyr Asn
                65                  70                  75
Lys Gln His Ile Ile Pro Leu Glu Asn Val Thr Ile Asp Ser Ile
                80                  85                  90
Lys Asp Glu Gly Asp Leu Arg Asn Gly Trp Leu Ile Lys Thr Pro
                95                 100                 105
```

```
Thr Lys Ser Phe Ala Val Tyr Ala Ala Thr Ala Thr Glu Lys Ser
                110                 115                 120
Glu Trp Met Asn His Ile Asn Lys Cys Val Thr Asp Leu Leu Ser
                125                 130                 135
Lys Ser Gly Lys Thr Pro Ser Asn Glu His Ala Ala Val Trp Val
                140                 145                 150
Pro Asp Ser Glu Ala Thr Val Cys Met Arg Cys Gln Lys Ala Lys
                155                 160                 165
Phe Thr Pro Val Asn Arg Arg His His Cys Arg Lys Cys Gly Phe
                170                 175                 180
Val Val Cys Gly Pro Cys Ser Glu Lys Arg Phe Leu Leu Pro Ser
                185                 190                 195
Gln Ser Ser Lys Pro Val Arg Ile Cys Asp Phe Cys Tyr Asp Leu
                200                 205                 210
Leu Ser Ala Gly Asp Met Ala Thr Cys Gln Pro Ala Arg Ser Asp
                215                 220                 225
Ser Tyr Ser Gln Ser Leu Lys Ser Pro Leu Asn Asp Met Ser Asp
                230                 235                 240
Asp Asp Asp Asp Asp Asp Ser Ser Asp
                245
<210> 45
<211> 247
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2790762CD1

<400> 45
Met Glu Thr Asp Glu Ser Pro Ser Pro Leu Pro Cys Gly Pro Ala
  1               5                  10                  15
Gly Glu Ala Val Met Glu Ser Arg Ala Arg Pro Phe Gln Ala Leu
                 20                  25                  30
Pro Arg Glu Gln Ser Pro Pro Pro Pro Leu Gln Thr Ser Ser Gly
                 35                  40                  45
Ala Glu Val Met Asp Val Gly Ser Gly Gly Asp Gly Gln Ser Glu
                 50                  55                  60
Leu Pro Ala Glu Asp Pro Phe Asn Phe Tyr Gly Ala Ser Leu Leu
                 65                  70                  75
Ser Lys Gly Ser Phe Ser Lys Gly Arg Leu Leu Ile Asp Pro Asn
                 80                  85                  90
Cys Ser Gly His Ser Pro Arg Thr Ala Arg His Ala Pro Ala Val
                 95                 100                 105
Arg Lys Phe Ser Pro Asp Leu Lys Leu Leu Lys Asp Val Lys Ile
                110                 115                 120
Ser Val Ser Phe Thr Glu Ser Cys Arg Ser Lys Asp Arg Lys Val
                125                 130                 135
Leu Tyr Thr Gly Ala Glu Arg Asp Val Arg Ala Glu Cys Gly Leu
                140                 145                 150
Leu Leu Ser Pro Val Ser Gly Asp Val His Ala Cys Pro Phe Gly
                155                 160                 165
Gly Ser Val Gly Asp Gly Val Gly Ile Gly Gly Glu Ser Ala Asp
                170                 175                 180
Lys Lys Asp Glu Glu Asn Glu Leu Asp Gln Glu Lys Arg Val Glu
                185                 190                 195
Tyr Ala Val Leu Asp Glu Leu Glu Asp Phe Thr Asp Asn Leu Glu
                200                 205                 210
Leu Asp Glu Glu Gly Ala Gly Gly Phe Thr Ala Lys Ala Ile Val
                215                 220                 225
Gln Arg Asp Arg Val Asp Glu Glu Ala Leu Asn Phe Pro Tyr Glu
                230                 235                 240
Val Cys Trp Gln Pro Leu Leu
                245
<210> 46
<211> 316
<212> PRT
```

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2869164CD1

<400> 46
```
Met Ala Glu Ala Ala Leu Glu Ala Val Arg Ser Glu Leu Arg Glu
 1               5                  10                  15
Phe Pro Ala Ala Ala Arg Glu Leu Cys Val Pro Leu Ala Val Pro
                20                  25                  30
Tyr Leu Asp Lys Pro Pro Thr Pro Leu His Phe Tyr Arg Asp Trp
                35                  40                  45
Val Cys Pro Asn Arg Pro Cys Ile Ile Arg Asn Ala Leu Gln His
                50                  55                  60
Trp Pro Ala Leu Gln Lys Trp Ser Leu Pro Tyr Phe Arg Ala Thr
                65                  70                  75
Val Gly Ser Thr Glu Val Ser Val Ala Val Thr Pro Asp Gly Tyr
                80                  85                  90
Ala Asp Ala Val Arg Gly Asp Arg Phe Met Met Pro Ala Glu Arg
                95                 100                 105
Arg Leu Pro Leu Ser Phe Val Leu Asp Val Leu Glu Gly Arg Ala
               110                 115                 120
Gln His Pro Gly Val Leu Tyr Val Gln Lys Gln Cys Ser Asn Leu
               125                 130                 135
Pro Ser Glu Leu Pro Gln Leu Leu Pro Asp Leu Glu Ser His Val
               140                 145                 150
Pro Trp Ala Ser Glu Ala Leu Gly Lys Met Pro Asp Ala Val Asn
               155                 160                 165
Phe Trp Leu Gly Glu Ala Ala Ala Val Thr Ser Leu His Lys Asp
               170                 175                 180
His Tyr Glu Asn Leu Tyr Cys Val Val Ser Gly Glu Lys His Phe
               185                 190                 195
Leu Phe His Pro Pro Ser Asp Arg Pro Phe Ile Pro Tyr Glu Leu
               200                 205                 210
Tyr Thr Pro Ala Thr Tyr Gln Leu Thr Glu Glu Gly Thr Phe Lys
               215                 220                 225
Val Val Asp Glu Glu Ala Met Glu Lys Val Pro Trp Ile Pro Leu
               230                 235                 240
Asp Pro Leu Ala Pro Asp Leu Ala Arg Tyr Pro Ser Tyr Ser Gln
               245                 250                 255
Ala Gln Ala Leu Arg Cys Thr Val Arg Ala Gly Glu Met Leu Tyr
               260                 265                 270
Leu Pro Ala Leu Trp Phe His His Val Gln Gln Ser Gln Gly Cys
               275                 280                 285
Ile Ala Val Asn Phe Trp Tyr Asp Met Glu Tyr Asp Leu Lys Tyr
               290                 295                 300
Ser Tyr Phe Gln Leu Leu Asp Ser Leu Thr Lys Ala Ser Gly Leu
               305                 310                 315
Asp
```

<210> 47
<211> 334
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3317629CD1

<400> 47
```
Met Thr Arg Ser Leu Phe Lys Gly Asn Phe Trp Ser Ala Asp Ile
 1               5                  10                  15
Leu Ser Thr Ile Gly Tyr Asp Asn Ile Ile Gln His Leu Asn Asn
                20                  25                  30
Gly Arg Lys Asn Cys Lys Glu Phe Glu Asp Phe Leu Lys Glu Arg
                35                  40                  45
```

```
Ala Ala Ile Glu Glu Arg Tyr Gly Lys Asp Leu Leu Asn Leu Ser
                50                  55                  60
Arg Lys Lys Pro Cys Gly Gln Ser Glu Ile Asn Thr Leu Lys Arg
                65                  70                  75
Ala Leu Glu Val Phe Lys Gln Gln Val Asp Asn Val Ala Gln Cys
                80                  85                  90
His Ile Gln Leu Ala Gln Ser Leu Arg Glu Glu Ala Arg Lys Met
                95                  100                 105
Glu Glu Phe Arg Glu Lys Gln Lys Leu Gln Arg Lys Lys Thr Glu
                110                 115                 120
Leu Ile Met Asp Ala Ile His Lys Gln Lys Ser Leu Gln Phe Lys
                125                 130                 135
Lys Thr Met Asp Ala Lys Lys Asn Tyr Glu Gln Lys Cys Arg Asp
                140                 145                 150
Lys Asp Glu Ala Glu Gln Ala Val Ser Arg Ser Ala Asn Leu Val
                155                 160                 165
Asn Pro Lys Gln Gln Glu Lys Leu Phe Val Lys Leu Ala Thr Ser
                170                 175                 180
Lys Thr Ala Val Glu Asp Ser Asp Lys Ala Tyr Met Leu His Ile
                185                 190                 195
Gly Thr Leu Asp Lys Val Arg Glu Glu Trp Gln Ser Glu His Ile
                200                 205                 210
Lys Ala Cys Glu Ala Phe Glu Ala Gln Glu Cys Glu Arg Ile Asn
                215                 220                 225
Phe Phe Arg Asn Ala Leu Trp Leu His Val Asn Gln Leu Ser Gln
                230                 235                 240
Gln Cys Val Thr Ser Asp Glu Met Tyr Glu Gln Val Arg Lys Ser
                245                 250                 255
Leu Glu Met Cys Ser Ile Gln Arg Asp Ile Glu Tyr Phe Val Asn
                260                 265                 270
Gln Arg Lys Thr Gly Gln Ile Pro Pro Ala Pro Ile Met Tyr Glu
                275                 280                 285
Asn Phe Tyr Ser Ser Gln Lys Asn Ala Val Pro Ala Gly Lys Ala
                290                 295                 300
Thr Gly Pro Asn Leu Ala Arg Arg Gly Pro Leu Pro Ile Pro Lys
                305                 310                 315
Ser Ser Pro Asp Asp Pro Asn Tyr Ser Leu Val Asp Asp Tyr Ser
                320                 325                 330
Leu Leu Tyr Gln

<210> 48
<211> 113
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3870488CD1

<400> 48
Met Asp Pro Lys Leu Leu Lys Gln Leu Arg Lys Ala Glu Lys Ala
 1               5                  10                  15
Glu Arg Glu Phe Arg Lys Lys Phe Lys Phe Glu Gly Glu Ile Val
                20                  25                  30
Val His Thr Lys Met Met Ile Asp Pro Asn Ala Lys Thr Arg Arg
                35                  40                  45
Gly Gly Gly Lys His Leu Gly Ile Arg Arg Gly Glu Ile Leu Glu
                50                  55                  60
Val Ile Glu Phe Thr Ser Asn Glu Glu Met Leu Cys Arg Asp Pro
                65                  70                  75
Lys Gly Lys Tyr Gly Tyr Val Pro Arg Thr Ala Leu Leu Pro Leu
                80                  85                  90
Glu Thr Glu Val Tyr Asp Asp Val Asp Phe Cys Asp Pro Leu Glu
                95                  100                 105
Asn Gln Pro Leu Pro Leu Gly Arg
                110
<210> 49
```

<211> 264
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3886318CD1

<400> 49
Met Leu Gly Ala Glu Thr Glu Glu Lys Leu Phe Asp Ala Pro Leu
1 5 10 15
Ser Ile Ser Lys Arg Glu Gln Leu Glu Gln Gln Val Pro Glu Asn
20 25 30
Tyr Phe Tyr Val Pro Asp Leu Gly Gln Val Pro Glu Ile Asp Val
35 40 45
Pro Ser Tyr Leu Pro Asp Leu Pro Gly Ile Ala Asn Asp Leu Met
50 55 60
Tyr Ile Ala Asp Leu Gly Pro Gly Ile Ala Pro Ser Ala Pro Gly
65 70 75
Thr Ile Pro Glu Leu Pro Thr Phe His Thr Glu Val Ala Glu Pro
80 85 90
Leu Lys Ala Asp Leu Gln Asp Gly Val Leu Thr Pro Pro Pro Pro
95 100 105
Pro Pro Pro Pro Pro Pro Ala Pro Glu Val Leu Ala Ser Ala Pro
110 115 120
Pro Leu Pro Pro Ser Thr Ala Ala Pro Val Gly Gln Gly Ala Arg
125 130 135
Gln Asp Asp Ser Ser Ser Ser Ala Ser Pro Ser Val Gln Gly Ala
140 145 150
Pro Arg Glu Val Val Asp Pro Ser Gly Gly Arg Ala Thr Leu Leu
155 160 165
Glu Ser Ile Arg Gln Ala Gly Gly Ile Gly Lys Ala Lys Leu Arg
170 175 180
Ser Met Lys Glu Arg Lys Leu Glu Lys Lys Gln Gln Lys Glu Gln
185 190 195
Glu Gln Val Arg Ala Thr Ser Gln Gly Gly His Leu Met Ser Asp
200 205 210
Leu Phe Asn Lys Leu Val Met Arg Arg Lys Gly Ile Ser Gly Lys
215 220 225
Gly Pro Gly Ala Gly Glu Gly Pro Gly Gly Ala Phe Ala Arg Val
230 235 240
Ser Asp Ser Ile Pro Pro Leu Pro Pro Pro Gln Gln Pro Gln Ala
245 250 255
Glu Glu Asp Glu Asp Asp Trp Glu Ser
260
<210> 50
<211> 185
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4043934CD1

<400> 50
Met Gly Gln Cys Leu Arg Tyr Gln Met His Trp Glu Asp Leu Glu
1 5 10 15
Glu Tyr Gln Ala Leu Thr Phe Leu Thr Arg Asn Glu Ile Leu Cys
20 25 30
Ile His Asp Thr Phe Leu Lys Leu Cys Pro Pro Gly Lys Tyr Tyr
35 40 45
Lys Glu Ala Thr Leu Thr Met Asp Gln Val Ser Ser Leu Pro Ala
50 55 60
Leu Arg Val Asn Pro Phe Arg Asp Arg Ile Cys Arg Val Phe Ser
65 70 75
His Lys Gly Met Phe Ser Phe Glu Asp Val Leu Gly Met Ala Ser
80 85 90

```
Val Phe Ser Glu Gln Ala Cys Pro Ser Leu Lys Ile Glu Tyr Ala
                 95                  100                 105
Phe Arg Ile Tyr Asp Phe Asn Glu Asn Gly Phe Ile Asp Glu Glu
                110                 115                 120
Asp Leu Gln Arg Ile Ile Leu Arg Leu Leu Asn Ser Asp Asp Met
                125                 130                 135
Ser Glu Asp Leu Leu Met Asp Leu Thr Asn His Val Leu Ser Glu
                140                 145                 150
Ser Asp Leu Asp Asn Asp Asn Met Leu Ser Phe Ser Glu Phe Glu
                155                 160                 165
His Ala Met Ala Lys Ser Pro Asp Phe Met Tyr Ser Phe Arg Ile
                170                 175                 180
Arg Phe Trp Gly Cys
                185
```

<210> 51
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4371445CD1

<400> 51

```
Met Phe Thr Ile Ile Phe Pro Val Cys Lys Asn Ser Met Pro Val
 1               5                   10                  15
Lys Lys Thr Asp Thr Asp Arg Ala Leu Ser Leu Leu Glu Glu Tyr
                20                  25                  30
Cys Lys Lys Leu Arg Lys Pro Glu Glu Gln Leu Leu Lys Asn Ala
                35                  40                  45
Val Lys Lys Val Met Gly Ile Phe Lys Ser Ser Leu Phe Gln Ala
                50                  55                  60
Leu Leu Gly Met Tyr Tyr Glu Ser Tyr Ser Ser Phe
                65                  70
```

<210> 52
<211> 434
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5527925CD1

<400> 52

```
Met Ala Ala Ala Ala Gly Ser Cys Ala Arg Val Ala Ala Trp Gly
 1               5                   10                  15
Gly Lys Leu Arg Arg Gly Leu Ala Val Ser Arg Gln Ala Val Arg
                20                  25                  30
Ser Pro Gly Pro Leu Ala Ala Ala Val Ala Gly Ala Ala Leu Ala
                35                  40                  45
Gly Ala Gly Ala Ala Trp His His Ser Arg Val Ser Val Ala Ala
                50                  55                  60
Arg Asp Gly Ser Phe Thr Val Ser Ala Gln Lys Asn Val Glu His
                65                  70                  75
Gly Ile Ile Tyr Ile Gly Lys Pro Ser Leu Arg Lys Gln Arg Phe
                80                  85                  90
Met Gln Phe Ser Ser Leu Glu His Glu Gly Glu Tyr Tyr Met Thr
                95                  100                 105
Pro Arg Asp Phe Leu Phe Ser Val Met Phe Glu Gln Met Glu Arg
                110                 115                 120
Lys Thr Ser Val Lys Lys Leu Thr Lys Lys Asp Ile Glu Asp Thr
                125                 130                 135
Leu Ser Gly Ile Gln Thr Ala Gly Cys Gly Ser Thr Phe Phe Arg
                140                 145                 150
Asp Leu Gly Asp Lys Gly Leu Ile Ser Tyr Thr Glu Tyr Leu Phe
                155                 160                 165
Leu Leu Thr Ile Leu Thr Lys Pro His Ser Gly Phe His Val Ala
```

170 175 180
Phe Lys Met Leu Asp Thr Asp Gly Asn Glu Met Ile Glu Lys Arg
185 190 195
Glu Phe Phe Lys Leu Gln Lys Ile Ile Ser Lys Gln Asp Asp Leu
200 205 210
Met Thr Val Lys Thr Asn Glu Thr Gly Tyr Gln Glu Ala Ile Val
215 220 225
Lys Glu Pro Glu Ile Asn Thr Thr Leu Gln Met Arg Phe Phe Gly
230 235 240
Lys Arg Gly Gln Arg Lys Leu His Tyr Lys Glu Phe Arg Arg Phe
245 250 255
Met Glu Asn Leu Gln Thr Glu Ile Gln Glu Met Glu Phe Leu Gln
260 265 270
Phe Ser Lys Gly Leu Ser Phe Met Arg Lys Glu Asp Phe Ala Glu
275 280 285
Trp Leu Leu Phe Phe Thr Asn Thr Glu Asn Lys Asp Ile Tyr Trp
290 295 300
Lys Asn Val Arg Glu Lys Leu Ser Ala Gly Glu Ser Ile Ser Leu
305 310 315
Asp Glu Phe Lys Ser Phe Cys His Phe Thr Thr His Leu Glu Asp
320 325 330
Phe Ala Ile Ala Met Gln Met Phe Ser Leu Ala His Arg Pro Val
335 340 345
Arg Leu Ala Glu Phe Lys Arg Ala Val Lys Val Ala Thr Gly Gln
350 355 360
Glu Leu Ser Asn Asn Ile Leu Asp Thr Val Phe Lys Ile Phe Asp
365 370 375
Leu Asp Gly Asp Glu Cys Leu Ser His Glu Glu Phe Leu Gly Val
380 385 390
Leu Lys Asn Arg Met His Arg Gly Leu Trp Val Pro Gln His Gln
395 400 405
Ser Ile Gln Glu Tyr Trp Lys Cys Val Lys Lys Glu Ser Ile Lys
410 415 420
Gly Val Lys Glu Val Trp Lys Gln Ala Gly Lys Gly Leu Phe
425 430

<210> 53
<211> 1629
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 129042CB1

<400> 53
gcgacctgta tgaggaggag gaggaggagg atgtgaagat ggcggacgtg cagatgctgc 60
tggaagagga aatcccgggg ggccgccggg ccctcttcga cagctacaca aatctggaac 120
gggtggccga ttactgcgag aacaactaca tacagtcagc agataagcag agagccctag 180
aagaaaccaa agcctacacc acccaatcct tagcaagtgt tgcctatctg ataaacacct 240
tggccaacaa tgtcctgcag atgctggata tccaggcatc ccagctacga aggatggaat 300
cttcaatcaa tcatatttca caaacagttg atattcataa agagaaagtt gcaagaagag 360
aaattggtat tttgactacc aataaaaaca cttcaaggac acataagatt attgctccag 420
ccaaccttga acgaccagtt cgttatatta gaaaacctat tgactataca attctagatg 480
atattggaca tggagtaaag gtgagtaccc agaacatgaa gatgggtggg ctgccgcgta 540
caacacctcc aactcagaag ccccctagtc cccctatgtc agggaaaggg acacttgggc 600
ggcactcccc ctatcgcaca ctggagccag tgcgtcctcc agtggtacca aatgattacg 660
tacctagccc aacccgtaat atggctccct cgcagcagag ccctgtgagg acagcttctg 720
tgaatcaaag aaatcgaact tacagcagca gtgggagtag tggagggagc cacccaagta 780
gtcggagcag cagtcgagag aacagtggaa gtggtagtgt ggggggttcct attgctgttc 840
ctactccatc tcctcccagt gtctttccag gtcatcctgt acagttctac agcatgaata 900
ggcctgcctc tcgccatact cccccaacaa taggggggctc gttgccctat agacgccctc 960
cttccattac ttcacaaaca agccttcaga atcagatgaa tggaggacct ttttatagcc 1020
agaatccagt ttcagataca ccacctccac cgccacctgt ggaagaacca gtctttgatg 1080
agtctccccc acctcctcct cctccagaag attacgaaga ggaggaagct gctgtggttg 1140
agtatagtga tccttatgct gaagaggacc caccgtgggc tccacgttct tacttggaaa 1200
aggttgtggc aatttatgac tatacaaaag acaaggaaga tgagctgtcc tttcaggaag 1260
gagccattat ttatgtcatc aagaagaatg acgatggttg gtatgaggga gttatgaatg 1320

gagtgactgg gctttttcct gggaattacg ttgagtctat catgcattat tctgagtaaa 1380
gctcagcagg gctgtgcttg cctcacagga atagtcaggt cttcccagat tatctgaagg 1440
ccctggggat tccactccag taaagtagaa tgaaggatac aaatgataaa aattacactt 1500
ttttttttgg tttattcccc agtattaaaa acaaagcaag ctgagtctga acaaatggat 1560
ctttctgcca tcatttgtac aatgctgagc tgtctggatt gaaataaaat gaccattttt 1620
atgtatgtc 1629

<210> 54
<211> 1257
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 778003CB1

<400> 54
gcacctgatc tttctcatcc ttccctgctc ttcccttcct ctccacctcc tcctcctcct 60
tggggaaagg ggcccggaga agggcatgtg ggggcccctc tgacagtggc ccgattgggg 120
tgacaggcgc ccaaatggcc aagtggctac gggactacct gagctttggg ggtcggaggc 180
cccctccgca gccgcccacc ccggactaca ccgagagcga catcctgagg gcctaccgcg 240
cgcagaagaa cctggacttt gaggacccct atgaggacgc ggagagccgc ttggagccgg 300
accccgcggg ccctggggac tccaagaacc ccggagatgc caagtatggt tctcccaagc 360
accggctcat caaggtggag gctgcggata tggccagagc caaggccctt ctgggcggcc 420
ccggggagga gctggaagcc gacactgagt atttagaccc ctttgatgct cagcctcatc 480
ctgcaccccc ggatgatggg tacatggagc cctacgatgc ccaatgggtc atgagtgaac 540
ttcccggcag aggggtgcag ctctatgaca ccccttatga ggaacaggac ccagagacag 600
cagatggacc cccttctggg cagaagcctc ggcagagccg gatgccccag gaagatgaac 660
ggccagcaga tgagtatgat cagccctggg agtggaagaa agaccacatc tccagggcgt 720
ttgcagtgca gtttgacagt ccagagtggg agaggactcc aggctcagcc aaggagctcc 780
ggagacctcc gcccagaagc ccccagcctg cggagcgtgt ggacccagcc ctgcccctgg 840
agaaacagcc gtggtttcat ggccccctga acagggcgga tgcagagagc ctcctgtccc 900
tctgcaagga aggcagctac ctagtgcggc tcagtgagac cagcccccag gactgctcct 960
tgtctctcag gagcagccag ggcttcctgc atctgaagtt cgcgcggacc cgtgagaacc 1020
aggtggtgct gggccaacac agcgggccct tccccagcgt gcccgagctc gtcctccact 1080
acagttcacg cccactgccg gtgcagggtg ccgagcatct ggctctgctg taccccgtgg 1140
tcacgcagac cccctgacag tgaccctcgg ccccctttttg agtcctcggg cccagaatcg 1200
tatcccaaag ccctcccatg gcctagaaaa taaataagtt attgttaaaa aaaaaaa 1257

<210> 55
<211> 1527
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1418671CB1

<400> 55
gcttcctggg cgccgtgggc gcggactgcg cgggctgcgc gggtgccgag gagcgcgagg 60
cgcggggaag gcgcacctgg ggtggccctg gcgtgcgggc ggcgacatgg aggacggcgt 120
gctcaaggag ggcttcctgg tcaagagggg ccacattgtc cacaactgga aggcgcgatg 180
gttcatcctt cggcagaaca cgctggtgta ctacaagctt gaggggggtc ggagagtgac 240
ccctcccaag ggccggatcc tcctggatgg ctgcaccatc acctgcccct gcctggagta 300
tgaaaaccga ccgctcctca ttaagctgaa gactcaaaca tccacggagt acttcctgga 360
ggcctgttct cgagaggagc gggatgcctg ggcctttgag atcaccgggg ctattcatgc 420
agggcagccg gggaaggtcc agcagctgca cagcctgaga aactccttca agctgccccc 480
gcacatcagc ctgcatcgca ttgtggacaa gatgcacgat agcaacaccg gaatccgttc 540
aagccccaac atggagcagg gaagcaccta taaaaagacc ttcctcggct cctccctggt 600
ggactggctc atctccaaca gcttcacggc cagccgtctg gaggcggtga ccctggcctc 660
catgctcatg gaggagaact tcctcaggcc tgtgggtgtc cgaagcatgg gagccattcg 720
ctctggggat ctggccgagc agttcctgga tgactccaca gccctgtaca cttttgctga 780
gagctacaaa aagaagataa gccccaagga agaaattagc ctgagcactg tggagttaag 840
tggcacggtg gtgaaacaag gctacctggc caagcaggga cacaagagga aaaactggaa 900
ggtgcgtcgc tttgttctaa ggaaggatcc agctttcctg cattactatg acccttccaa 960
agaagagaac aggccagtgg gtgggttttc tcttcgtggt tcactcgtgt ctgctctgga 1020
agataatggc gttcccactg ggttaaagg gaatgtccag ggaaacctct tcaaagtgat 1080

tactaaggat gacacacact attacattca ggccagcagc aaggctgagc gagccgagtg 1140
gattgaagct atcaaaaagc taacatgaca aggacctgag ggaaccagga ttcctccctc 1200
ctaccagatg acacagacaa gagttcctgg agaatgggag tgttaagact tttgacttct 1260
ttgtaagttt tgtactgctt tggagagtga atgctgccaa gagttcctca gattacaaac 1320
agcagtggtg ccatttcctt ccccatcttc atgttacaaa cctggaaagg ctagaacagc 1380
cattaggcgt cagcatcttg acttttcccc agcatcacaa acagccattt cctcgggcac 1440
caaagtaggt tccctttgtt ggaacaatta cactggccat gccataatgt tgaataaaac 1500
tctcttctta tgaaaaaaaa aaaaaaa 1527

<210> 56
<211> 2220
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1456841CB1

<400> 56
ctgctgtcct tccaccacca gcaccggacc acctgctcca agaccagcct cctgggggga 60
ccacgcaccc ggccttcact ggcacccagg gagccgtcct cagcagcgtc aacatgtcaa 120
ggcccagcag cagagccatt tacttgcacc ggaaggagta ctcccagaac ctcacctcag 180
agcccaccct cctgcagcac agggtggagc acttgatgac atgcaagcag gggagtcaga 240
gagtccaggg gcccgaggat gccttgcaga agctgttcga gatggatgca cagggccggg 300
tgtggagcca agacttgatc ctgcaggtca gggacggctg gctgcagctg ctggacattg 360
agaccaagga ggagctggac tcttaccgcc tagacagcat ccaggccatg aatgtggcgc 420
tcaacacatg ttcctacaac tccatcctgt ccatcaccgt gcaggagccg ggcctgccag 480
gcactagcac tctgctcttc cagtgccagg aagtgggggc agagcgactg aagaccagcc 540
tgcagaaggc tctggaggaa gagctggagc aaagacctcg acttggaggc cttcagccaa 600
gccaggacag atggaggggg cctgctatgg aaaggccgct ccctatggag caggcacgct 660
atctggagcc ggggatccct ccagaacagc cccaccagag gaccctagag cacagcctcc 720
caccatcccc aaggcccctg ccacgccaca ccagtgcccg agaaccaagt gcctttactc 780
tgcctcctcc aaggcggtcc tcttccccccg aggacccaga gagggacgag gaagtgctga 840
accatgtcct aagggacatt gagctgttca tgggaaagct ggagaaggcc caggcaaaga 900
ccagcaggaa gaagaaattt gggaaaaaaa acaaggacca gggaggtctc acccaggcac 960
agtacattga ctgcttccag aagatcaagt acagcttcaa cctcctggga aggctggcca 1020
cctggctgaa ggagacaagt gcccctgagc tcgtacacat cctcttcaag tccctgaact 1080
tcatcctggc caggtgccct gaggctggcc tagcagccca agtgatctca cccctcctca 1140
cccctaaagc tatcaacctg ctacagtcct gtctaagccc acctgagagt aacctttgga 1200
tggggttggg cccagcctgg accactagcc gggccgactg gacaggcgat gagcccctgc 1260
cctaccaacc cacattctcg gatgactggc aacttccaga gccctccagc caagcaccct 1320
taggatacca ggaccctgtt tcccttcggc ggggaagtca taggttaggg agcacctcac 1380
actttcctca ggagaagaca cacaaccatg accctcagcc tggggacccc aactccaggc 1440
cctccagccc caaacctgcc cagccagccc tgaaaatgca agtcttgtac gagtttgaag 1500
ctaggaaccc acgggaactg actgtggtcc agggagagaa gctggaggtt ctggaccaca 1560
gcaagcggtg gtggctggtg aagaatgagg cgggacggag cggctacatt ccaagcaaca 1620
tcctggagcc cctacagccg gggacccctg ggacccaggg ccagtcaccc tctcgggttc 1680
caatgcttcg acttagctcg aggcctgaag aggtcacaga ctggctgcag gcagagaact 1740
tctccactgc cacggtgagg acacttgggt ccctgacggg gagccagcta cttcgcataa 1800
gacctgggga gctacagatg ctatgtccac aggaggcccc acgaatcctg tcccggctgg 1860
aggctgtcag aaggatgctg ggataagcc cttaggcacc agcttagaca cctccaagaa 1920
ccaggccccg ctgatgcaag atggcagatc tgatacccat tagagccccg agaattcctc 1980
ttctggatcc cagtttgcag caaaccccac accccagctc acacagcaaa aacaatggac 2040
aggcccagag ggtgaagcaa acagtgtccc ttctggctgt gttggagcct ccccagtaac 2100
cacctattta ttttacctct ttcccaaacc tggagcattt atgcctaggc ttgtcaagaa 2160
tctgttcagt ccctctcctt ctcaataaaa gcatcttcaa gcttgtaaaa aaaaaaaaaa 2220

<210> 57
<211> 2895
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2020010CB1

<400> 57

ccaggcccga agccgaggcg gggccgggat gcggcgctga ggcccagcat ggccggcccg 60
ggccccacct tcccgctgca ccggctcgtc tgggcgaacc ggcatcgcga actggaggcc 120
gcactgcaca gccaccagca cgacattgaa caggaggacc cccgcgggcg gaccccactg 180
gagctggccg tgtctctggg aaacctggag tctgtgagag tgctccttcg acacaatgcc 240
aacgtgggca aagagaaccg ccagggctgg gcagtcctgc aggaggcagt cagcactgga 300
gaccccgaga tggtgcagct ggtgctccag tatcgggact accagagggc cacgcagagg 360
ctggcgggca ttccggaact gctcaacaaa cttcgccagg cccccgattt ctacgttgag 420
atgaagtggg agttcaccag ctgggtgccc cttgtgtcta agatgtgccc aagcgatgtg 480
taccgcgtgt ggaagcgggg tgagagcctg cgagtagaca ccagtctcct gggcttcgag 540
cacatgacct ggcagcgggg ccggaggagc ttcatcttca agggccagga ggcaggagcc 600
ctggtgatgg aagtggacca tgaccggcag gtggtgcatg tggagacact ggggctcact 660
ctgcaggagc ccgaaacact gctggccgcc atgcggccca gcgaggagca tgtggccagt 720
cgcctcacct ctcctatcgt ctccacccac ctggacactc gtaatgtggc ctttgagagg 780
aacaaatgtg gtatctgggg ctggcggtct gagaagatgg aaactgttag cggctacgag 840
gccaaggtgt acagtgccac caacgtggag ctggtgacac gcacacgcac ggagcacctc 900
tctgatcagg acaagtcgag gagcaaagcg ggaagactc cattccagtc cttcctgggg 960
atggcgcagc agcattcctc ccacaccggg gcccccgtgc agcaggcagc cagcccccacc 1020
aaccccacag ccatctcccc tgaggagtac ttcgacccca acttcagcct ggagtcacgg 1080
aacattggcc gccccatcga gatgtccagc aaagtacaga ggtgaggtct gagagctggc 1140
tggggacttg cctcgggaca agggctcttg cagacccctc tctgggcctg tcatagttag 1200
ggacccactc cctcggctgg cttctctctg gactccactc ctggagggca ggagtcatgt 1260
ctgccttatc catggtcctc agcccttagc aaggggcttg gcacagaaaa ccggtagttt 1320
gtctttgatg aatggatggt gccacccatg tatggttttc tattgaattt catgagtacc 1380
tgctggggcc agcgtggcac agtgggaagg gcccccagcg acgtggcctg ggagtggagc 1440
gtctgccttg gtggaacaga acagttacgc tcttgcctcg tgaggagcct tgagcaggta 1500
taggcggtaa cagcaggcac agtgcctgca caagcctagc gctctcccca ccatctcagg 1560
aggattctct tagccaccca acagtcgctg ggattcgaac cctggcagtc ttgccctggg 1620
gtgtgctgga cataaaagtt tggcaacacg tgagctggtg accagctctg ggctgaggag 1680
gaaaacgggg ctgtgggcca ggcccagaga gaagcccaca gcctggcacc tggcctcctt 1740
gtccagacca gatgagctgg tacgaaatcc tcaggagccc cagcctgggc ccagggagga 1800
gggcagcttg ggccacgtgg ccaggacacc agctcccggg ggaggcgggc agcggcatct 1860
gagcagagca gggcactcca ggccaggcag aaggtgggta aaggcagctg cccacgaacc 1920
agagggcagt cctcaatgga agggccacca gccgtgcctc acccatgtcc tgtggtcggc 1980
tgggcaggtt caaggcaaca ctgtggctga gtgaagagca cccgctctcc ctgggtgacc 2040
aggtgacccc catcatcgac ctaatggcca tcagcaacgc tcactttgcc aagctgcgcg 2100
acttcatcac tctgcgcctt ccacctggct tccccgtcaa aattgagatt cccctttttcc 2160
acgtgctcaa tgcccgcatc accttcagca acctgtgtgg ctgtgatgag cccctgagct 2220
ccgtgtgggt gccggccccc agctctgctg tcgccgcatc agggaaccct ttcccgtgcg 2280
aggtggaccc caccgtgttt gaagtgccca acgggtacag cgtgctgggc atggagcgca 2340
acgagcccct ccgggacgag gacgatgacc tcctgcagtt cgccatccag cagagcctgc 2400
ttgaagcggg cactgaggcg gagcaggtga ccgtctggga agccctgacc aacacccggc 2460
ccggtgcccg ccctcctccc caggccacgg tttatgagga acagcttcag ctggagcggg 2520
ccctccagga aagcctgcag ctgtccacag agcccagggg cccaggatcc cctcccagga 2580
cacccccagc ccccggtcca cccagctttg aagagcagct gcgcctggcc ctggagttgt 2640
cttcacggga gcaggaggag cgggagcggc gcgggcagca ggaggaggag gacttacagc 2700
ggatcctgca gctgtcactc actgagcact gagccatagc cccgggaggg ctggccaggc 2760
cactccctgc ccgcttttgt aatttattta tttataaact ctctgctgct gagcttgggg 2820
cctggagccc caggaatgag caggcagggg agactgagat ggaaataaag agactgtcgc 2880
agcaaaaaaa aaaaa 2895

<210> 58
<211> 2801
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2149037CB1

<400> 58
gcgccgctga gcgctgactg ggtgcgagtg gggaagctgc taacccgacc cggattggcg 60
ctgaggtggc ccgtggggca gggcagatga ttctggacca gatgaagcct gaggagcctt 120
ccagctctaa gatagcagga taggagactt ctaagattgg agctgcagaa gacttgccag 180
cccaccagca caatgtcagg aagccataca cctgcctgtg gcccttttctc agccctgact 240
ccgagcatat ggccccagga gatcttggcc aagtacacgc agaaggaaga gtcagcagag 300
caaccagagt tctactacga tgagtttggt ttccgtgtgt acaaggaaga aggtgatgag 360
cctggctcca gtctgctggc gaactcccct ctgatggagg atgctccaca gaggctgcgg 420

tggcaggccc acctggagtt cacccataac cacgatgtgg gggatctcac ctgggacaag 480
attgccgtct ccctaccccg ctctgagaag ctccgctccc tggtgctggc cggcatccca 540
catggcatga ggccacagct gtggatgcgg ctctctgggg ccctgcagaa gaagaggaac 600
tctgagctgt cctaccgcga gattgtgaag aacagctcca acgatgagac catcgctgcc 660
aagcagatcg agaaggacct gctccgcacc atgcccagca acgcctgctt cgccagcatg 720
ggtagcatcg ggtgccccg cctgcgcagg gtgctccggg ccctggcctg gctctaccca 780
gagatcggct actgccaggg caccggcatg gtggccgcct gcctcctgct gttcctggag 840
gaggaggacg ccttctggat gatgtctgcc atcatcgagg acctgctccc cgcctcctac 900
ttcagcacca ccctgctggg tgtccagact gaccagcggg tcctgcgcca cctcattgtc 960
cagtacctgc ctcgcctgga caagctgctc caggagcatg acattgagct gtccctgatc 1020
acactgcact ggttcctcac ggccttcgcc agcgtggtgg acatcaagct gctcctgcgc 1080
atctgggacc tgtttttcta cgagggctcc cgggtgctgt tccagctcac gctgggcatg 1140
ctgcacctca aggaggaaga gctgatccag tcagagaact cggcctccat cttcaacacg 1200
ctatcggata tcccgtcgca gatggaggac gcggagctgc ttctgggggt ggccatgcgg 1260
ctggccggct ccctcaccga tgtggccgtg gagactcagc gccgcaagca cctggcctat 1320
ctcattgcag accagggcca gctcctgggg gccggcaccc tcaccaacct ctctcaggtt 1380
gttcgccgca ggacccagcg gaggaagtcc accatcactg ctctgctctt cggggaggat 1440
gacctggagg cactcaaggc caagaacatc aagcagacgg aactggtggc tgacctccgg 1500
gaagccatcc tgcgcgtggc acgccacttc cagtgcacag accccaaaaa ctgcagcgtg 1560
gagctgactc cagactatag catggagagc caccagcggg accacgagaa ctacgtggcg 1620
tgctcacgca gccaccggcg ccgagccaag gccctgctgg actttgagcg gcacgacgac 1680
gacgagctgg gcttccgcaa gaacgacatc atcacaatcg tgtctcagaa ggacgagcac 1740
tgctgggtgg gggagctcaa cggcctgcga ggctggtttc cagccaagtt cgtggaagtc 1800
ctggatgagc gcagcaaaga gtactccatc gcgggggatg actcggtgac ggagggggtc 1860
acagacctcg tgcgagggac cctctgcccg gcccttaagg ccctgttcga acatggactg 1920
aagaagccat ccctgcttgg gggcgcctgc cacccctggc tgtttatcga ggaggctgca 1980
ggccgggagg tcgagagaga ctttgcctcc gtgtattccc gtctggtgct ctgtaagacc 2040
ttcaggttgg atgaagatgg caaagtcctg accccggagg agctgctcta ccgggctgtg 2100
cagtctgtga acgtgaccca cgatgcagtg catgcacaaa tggatgtgaa gctccgctca 2160
ctgatctgcg tggggctcaa tgagcaggtg ctgcacctgt ggctggaggt gctctgctcc 2220
agcctgccca ccgtggagaa gtggtaccag ccctggtcct tcctgcgcag cccgggctgg 2280
gtccagatca agtgtgagct ccgagtcctc tgctgctttg ccttcagcct ctcccaggac 2340
tgggagctcc ctgcgaagag agaggcgcag cagcccctga aggagggcgt ccgggacatg 2400
ctggtgaagc accacctctt cagctgggat gtggacgggt gacccccctcc tccccagccc 2460
aacctcgggc ctgcgtctga ggtggcccag gaccccaagc tgcagagccc agggaagagc 2520
agctccagag ccctggccgg ggccgcggga tatcaatatc aggctgcccc actccacgtt 2580
ccccagcaca tcccaggtgg tgggagcaga gggtaccctg ccccaccagg gtccttaggg 2640
atgctctagg ccaaaccaca gtttgtacca aaaccttgt gaggaggtgg gggagccatg 2700
tctgtgctca ggaagaggga aggggatggg ggtggctagt aggctcctgg cctctttggt 2760
ttataaataa actgtgtctg tctttgagaa aaaaaaaaaa a 2801

<210> 59
<211> 599
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2162179CB1

<400> 59
aggttatcgt taggcatctc ccaggcgacc ggctccgcag caagatggcg gacgagaagg 60
acagggaaga gataatagta gcagaatttc acaaaaaaat caaagaggca tttgaagtct 120
ttgaccatga gtcgaataat acagtggatg tgagagagat tggaacaatt atcaggtcat 180
taggatgctg tcctacggaa ggagagctgc atgatctgat tgcagaggta gaggaagaag 240
aacccactgg atacattcga ttcgaaaaat ttcttccggt gatgacagaa atactactag 300
aaagaaaata cagaccaatt ccagaagatg tccttcttcg agcttttgag gttttagatt 360
cagctaaacg tgggtttctt actaaggacg agctgatcaa gtatatgact gaagaaggta 420
agtgtgattt attacttatc acaatgactt atgtgaggaa ttaataattt gttaacagtt 480
atgcgaaagt tatagggat actttaaaat cagtcattct ggtgaaagtt atttactgtt 540
ccagcctggg cgacagagca agactccatc tcacaaaaaa aaaaaaaggg gggggaggg 599

<210> 60
<211> 2065
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2244706CB1

<220>
<221> unsure
<222> 2060-2061
<223> a, t, c, g, or other

<400> 60
gtaattactg gaaccacaga aaattcacct gcagatcggt gcaagaaaat ccatgctggc 60
gatgaagtga ttcaagttaa tcatcagact gtgcctctta tacctagaag tcccacaagc 120
agcgttgcca cgccttccag caccatcagt acacccacca aaagagacag ttctgccctc 180
caggatctct acattccccc tcctcctgca gaaccatata ttcccaggga tgaaaaagga 240
aaccttcctt gtgaagacct cagaggacat atggtgggca agccagtgca taagggatct 300
gaatcaccaa attcatttct ggatcaggaa tatcgaaaga gatttaatat tgtcgaagaa 360
gatactgtct tatattgcta tgaatatgaa aaaggaagat caagtagtca aggaagacga 420
gaaagcaccc caacttatgg caagctacga cctatatcta tgccagtgga atataattgg 480
gtgggggact atgaagatcc aaataagatg aagagagata gtagaagaga aaactctcta 540
cttcggtata tgagcaatga aaagattgct caagaagaat acatgtttca gagaaacagc 600
aaaaaggaca cagggaagaa gtcaaaaaag aagggtgata agagtaatag cccaactcac 660
tattcattgc tacctagttt acaaatggat gcactgagac aagacatcat gggcactcct 720
gtgccagaga ccacactata ccatacattt cagcagtcct cactgcagca caaatcaaag 780
aagaaaaaca aaggtcctat agcaggcaag agcaaaagac gaatttcttg caaagatctt 840
ggccgtggtg actgtgaggg ctggctttgg aaaaagaaag atgcgaagag ttactttca 900
cagaaatgga aaaaatattg gtttgtccta aaggatgcat ccctttattg gtatattaat 960
gaggaggatg aaaaagcaga aggattcatt agcctgcctg aatttaaaat tgatagagcc 1020
agtgaatgcc gcaaaaaata tgcattcaaa gcctgtcatc ctaaaatcaa aagcttttat 1080
tttgctgctg aacatcttga tgatatgaac aggtggctta acagaattaa tatgctgact 1140
gcaggatatg cagaaagaga gaggattaag caggaacaag attactggag tgagagtgac 1200
aaggaagaag cagatactcc atcaacacca aaacaagata gccctccacc cccatatgat 1260
acatacccac gacctccctc gatgagttgc gccagtcctt atgtggaagc aaaacatagc 1320
cgactttcct ccacggagac ttctcagtct cagtcttctc atgaggagtt tcgccaggaa 1380
gtaactggga gcagtgcagt gtctcccatt cgcaagacag ccagtcagcg ccgctcctgg 1440
caggatttaa ttgagacgcc actgacaagt tcaggcttac actatcttca gactctgccc 1500
ctggaggatt ctgtcttctc tgactccgcg gccatctccc cagagcacag gcggcagtct 1560
accctgccaa ctcagaaatg ccacctgcag gatcactatg ggccataccc cttagctgag 1620
agtgagatga tgcaagtgct aaatggaaat gggggcaagc ctcgaaggtt tactctgcct 1680
cgagatagcg ggttcaacca ttgctgtctg aatgctccag ttagtgcctg tgacccacag 1740
gatgacgtgc aaccaccaga ggtggaggaa gaggaggacg atgaggagga agcatgggag 1800
gcagccggtg gaaacatggg agaaaaaagc ctattcactg cgagagtggg tagacctttc 1860
atgcaaaacg gatccactct gtggcactaa ccattggact acagattata ccaatgttag 1920
agattagatt ggaaattggg gtgcggattc ggttaaccta tggtacctgg ctctaggggg 1980
gccgtccaat tgcctcagta gtgtttcgcg gcccggcgtg tttaaagtgt atgggaactg 2040
ggtcccatta gcctgagcan ncccc 2065

<210> 61
<211> 2330
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2316805CB1

<400> 61
gctttgcaga gtgattatca gcacagttcc ctgccctgga taaggaacag ctacagtcgc 60
tgttaaatgt gcctgaaaag caatttgcaa tctttgcatt aggatttcag atgcatgcca 120
ggtttccact gattgccaga actcgagatc actacacatg gatccccaaa atcaacatgg 180
cagtggcagt tcgttagttg tgatccagca gccttctttg gatagccgtc agagattaga 240
ctatgagaga gagattcagc ctactgctat tttgtcctta gaccagatca aggccataag 300
aggcagcaat gaatacacag aagggccttc ggtggtgaaa agacctgctc ctcggacagc 360
accaagacaa gaaaagcatg aaaggactca tgaaatcata ccaattaatg tgaataataa 420
ctacgagcac agacacacaa gccacctggg acatgcagta ctcccaagta atgccagggg 480
ccccatttg agcagatcaa ccagcactgg aagtgcagcc agctctggga gcaacagcag 540
tgcctcttct gaacagggac tgttaggaag gtcaccacca accagaccag tccctggtca 600
taggtctgaa agggcaatcc ggacccagcc caagcaactg attgtggatg acttgaaggg 660

ttccttgaaa gaggacctga cacagcacaa gttcatttgt gaacagtgtg ggaagtgcaa 720
gtgtggagaa tgcactgctc ccaggaccct accatcctgt ttggcctgta accggcagtg 780
cc~~tct~~gctct gctgagagca tggtggaata tggaacctgc atgtgcttag tcaagggcat 840
cttctaccac tgctccaatg acgacgaagg ggattcctat tcagataatc cttgctcctg 900
ttcacaatca cactgctgct ctagatacct gtgtatggga gccatgtctt tatttttacc 960
ttgcttactc tgttatcctc ctgctaaagg atgcctgaag ctgtgcagga ggtgttatga 1020
ctggatccat cgcccagggt gcagatgtaa gaactccaac actgtctatt gtaagctgga 1080
gagctgcccc tcccggggtc agggtaaacc atcatgattt ttggaggtgg gttgtacctc 1140
ctgaacttct agctttcaag ttgtggctgt tttttgtttt tgtttttgtt tttgttttct 1200
ttagaatttt tccctgtttc ccaccttctc ttcccctgtt gccaaggtct aactcatgga 1260
tttttctctt tcctcatgga tgatcttcag caagagtgga ctgggaagct gcacctggct 1320
cccactttca acaagagcct ctgccatcca cttgagggta ttgagagcca gtgggctttt 1380
gtgtagcctt tttgttctgc aagcaacttt ctaaagttgt gtacatgaac atacacccac 1440
atccagacta cagtgattta gagttgtttt gattgggtac cgtgggagca gggaaattgg 1500
ttttttaaaa agcaactgtt taattgctta aataagctat gtattaaatc tgtctccagt 1560
tagggctatc ttcctagcat aggcccctta agtagcatgg gggatatatt ttttgctata 1620
acgtaaaaat tttcctttaa ccactgccct ctccttcttt ctccttcaag gttctttccc 1680
cctcagtttt gttgttgtct tactctggag atgccaagtg tatttttttct ttctatgtaa 1740
ttttagattc gccttacaat gtaaatcttc acattggaga taatattggt tggaccttgc 1800
ccatcttcac tctagccttc gtatttgtga aggactcagc caccttcctt cttcacccca 1860
tgcttctcac caaatttttg ttgtcattga gggcacttgg ataactcaag ttgatattta 1920
tagctgatca atctatatgt gtcacagaac tatgctgcct aaagtgatct tggctcctta 1980
atggtccttt tggccccttg gatagttaac agctgagtaa ttctaatctc ttctgtgttt 2040
tccttgcctt aaccacaaat tgtggtgctt tttgtatatt ttatgtataa atcacaaagt 2100
tgaattctga ctatttttaa gacaaaagtc tgttaaactt ttttattgta aagaatattt 2160
attatgcgaa tctctattat tttatggtat ttattgcaaa agactgttga aatgtactca 2220
tgtttgaata taacaaaata tcaatactta acggaaaata aggtgacacg aagaaagtac 2280
atatgttaac tataatgcag aaaatatatt aattaatgaa aaaaaaaaaa 2330

<210> 62
<211> 2610
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2320010CB1

<400> 62
aagtatgctt gtttcacttt agatatatgg ggaaaagaaa catagcaagg gtgcatgatg 60
cctggctgtc aaaacacttc ggaatagacc gaaaatcgca aaccatgcct gctcttcgaa 120
acagatcagg agtaatgcag gcccggcttc agcatcttag tagcctagaa agttcattta 180
cacttaatca cagttctaca acaactgaag cagacatttt ccaccaggca cttcttgccg 240
cgaatacagc tactgaagtt tccctaacag tactagacac catatcattt ttcactcagt 300
gcttcaagac ccaactttta aataatgatg gccataaccc attaatgaaa aaagtgtttg 360
atatacatct tgcttttctt aaaaatggac aatctgaagt gtcgctgaaa catgtatttg 420
cctcactgag agctttcatc agtaagtttc cttcagcatt tttcaaagga agagtaaaca 480
tgtgtgctgc attttgctat gaggttttaa agtgctgcac atcgaagatt agctcaacca 540
ggaatgaagc atctgcactt ttgtatcttt tgatgagaaa caactttgag tataccaaaa 600
ggaaaacctt tttgaggaca catctacaga taataattgc tgtaagccaa ctgatagctg 660
atgtagcact aagcggagga tcaagatttc aggagtcttt attcattatc aataattttg 720
caaatagtga cagacctatg aaggcaactg cctttcccgc agaagtcaaa gacttgacca 780
agagaatccg cactgttctt atggccactg cccaaatgaa ggagcatgag aaagaccctg 840
aaatgctaat tgatctccag tatagcttag ccaagtccta tgcaagcacc ccagagctca 900
ggaaaacctg gcttgatagc atggccaaga ttcatgtaaa aaatggagat ttttcagagg 960
ctgcgatgtg ttatgtccat gtagcagctc tagttgcaga gtttcttcat cgaaaaaaat 1020
tatttcctaa cggatgttca gcgttcaaga aaattactcc caatatagat gaagaaggag 1080
caatgaaaga agatgctggg atgatggatg tccattatag tgaagaggtc ttgctggagt 1140
tgctagaaca atgtgtggat ggcttatgga aggcagaacg ttatgaaata atttctgaga 1200
tttccaagtt gatcgttcca atttatgaga aacgtcgtga gtttgagaaa cttactcaag 1260
tttatagaac tcttcatgga gcttacacaa aaattctgga agttatgcat acaaaaaaga 1320
gacttttagg cactttcttc agagttgcct tttatggcca atcttttttt gaagaagaag 1380
atggaaagga gtacatctat aaagaaccaa agctcactgg cctctcagaa atttccttga 1440
gacttgttaa actttatggt gaaaagtttg gtacggagaa tgtcaaaata attcaggatt 1500
cagacaaggt aaatgccaaa gagcttgatc caaaatatgc tcatatacaa gttacttatg 1560
tgaagcctta ctttgatgac aaagaactca cagaaaggaa gaccgagttt gaaagaaatc 1620
ataatatcag cagatttgtt tttgaggccc cttacacttt atcaggcaaa aaacagggct 1680

gtatagaaga acagtgcaaa cgccgtacaa tcttgacaac ttcaaactcg tttccttacg 1740
tgaagaagag gattcctatt aactgtgaac agcagattaa tttaaaacca attgatgttg 1800
ccactgatga aataaaagat aaaactgcag agctgcaaaa gctttgctcc tctactgacg 1860
tggacatgat tcagctccaa cttaaattgc agggctgtgt ttctgtgcag gtcaatgctg 1920
gtccattagc atatgcaaga gctttcttaa atgacagcca agctagcaag tatccaccta 1980
agaaagtgag tgagttgaaa gacatgttta ggaaatttat acaagcatgc agcattgcac 2040
ttgaactaaa tgagcggcta attaaagaag atcaagttga gtaccatgaa gggctaaagt 2100
caaatttcag agacatggta aaagaattat ctgacattat ccatgagcag atattacaag 2160
aagacacaat gcattctccc tggatgagca acacattaca tgtattttgt gcaattagtg 2220
gtacatcaag tgaccgaggt tatggttccc caagatacgc tgaagtgtga ggcaatgcag 2280
atgtacgtga caatgagact gacctttctc aggaatattt ggagctgtgc aaatgttaaa 2340
atttaaagat ttgatataca tggagtgttt cttctcgaca ccaaaatttt catgtgttcc 2400
agcagggtgc ttacatattt gtaaataagc aacttgaaag tgcctggaaa attgcaccac 2460
tgtgtttggt ttgtactttt ttaggtaaat ctatatgctg aaaagtagag ctcaaaaaca 2520
gtagttcaat ttgcttaatt attgcttaaa ataatggtac tatgtaaaat tgtataatgg 2580
aatacaataa aaggtaaacc ttaaaacaaa 2610

<210> 63
<211> 1035
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2564901CB1

<400> 63
agtccacctt gcgaccgtat ccgctagcgc ggcctgggat gcgcttgggc tccctgttcg 60
ttcccacatg cagggcagca caaggagaat gggcgtcatg actgatgtcc accggcgctt 120
cctccagttg ctgatgaccc atggcgtgct agaggaatgg gacgtgaagc gcttgcagac 180
gcactgctac aaggtccatg accgcaatgc caccgtagat aagttggagg acttcatcaa 240
caacattaac agtgtcttgg agtccttgta tattgagata aagagaggag tcacggaaga 300
tgatgggaga cccatttatg cgttggtgaa tcttgctaca acttcaattt ccaaaatggc 360
tacggatttt gcagagaatg aactggattt gtttagaaag gctctggaac tgattattga 420
ctcagaaacc ggctttgcgt cttccacaaa catattgaac ctggttgatc aacttaaagg 480
caagaagatg aggaagaagg aagcggagca ggtgctgcag aagtttgttc aaaacaagtg 540
gctgattgag aaggaagggg agttcaccct gcacggccgg gccatcctgg agatggaaca 600
atacatccgg gagacgtacc ccgacgcggt gaagatctgc aatatctgtc acagcctcct 660
catccagggt caaagctgcg aaacctgtgg gatcaggatg cacttaccct gcgtggccaa 720
gtacttccag tcgaatgctg aaccgcgctg cccccactgc aacgactact ggccccacga 780
gatcccaaaa gtcttcgacc ctgagaagga gagggagtct ggtgtcttga aatcgaacaa 840
aaagtccctg cggtccaggc agcattagcc atcgtgccct gctgaggggc tggctgcctt 900
gagtggcctg atcgccacag cccttcttgg aagaaaggcg tctgtgtttc aggttccacg 960
cgagtcacct ctttcgtctt aatgttcacc gtccacagct ttggaataaa ccatcctggg 1020
aaaaaaaaaa aaaaa 1035

<210> 64
<211> 1838
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2615168CB1

<400> 64
tgagtggagt tcactcacat ggattgaggc ccagttcctg ggagaagaga tgctgggcag 60
gaaggtgtct gcatgtggga ctctgtacag cccggtcctc tcccacatct gggaggggcc 120
agagtcagac aactgctggg ttcgtcccta agagaggtca tctgactggc tgttcagcct 180
aggctgcaca cacccccact ttcctctacc aggccacacc ggaggcagtg ctcacacagg 240
caagctacca ggccacaaca acgacaccca cctcacctct ggcacctctg agcatccacg 300
tacttgcaag aactcttgct cacatcagct aagagattgc acctgctgac ctagagattc 360
cggcctgtgc tcctgtgctg ctgagcaggg caaccagtag caccatgtct gtgactggcg 420
ggaagatggc accgtccctc acccaggaga tcctcagcca cctgggcctg gccagcaaga 480
ctgcagcgtg gggaccctg ggcaccctca ggaccttctt gaacttcagc gtggacaagg 540
atgcgcagag gctactgagg gccattactg gccaaggcgt ggaccgcagt gccattgtgg 600
acgtgctgac caaccggagc agagagcaaa ggcagctcat ctcacgaaac ttccaggagc 660

gcacccaaca ggacctgatg aagtctctac aggcagcact ttccggcaac ctggagagga 720
ttgtgatggc tctgctgcag cccacagccc agtttgacgc ccaggaattg aggacagctc 780
tgaaggcctc agattctgct gtggacgtgg ccattgaaat tcttgccact cgaaccccac 840
cccagctgca ggagtgcctg gcagtctaca aacacaattt ccaggtggag gctgtggatg 900
acatcacatc tgagaccagt ggcatcttgc aggacctgct gttggccctg gccaaggggg 960
gccgtgacag ctactctgga atcattgact ataatctggc agaacaagat gtccaggccc 1020
tgcagcgggc agaaggacct agcagagagg aaacatgggt cccagtcttc acccagcgaa 1080
atcctgaaca cctcatccga gtgtttgatc agtaccagcg gagcactggg caagagctgg 1140
aggaggctgt ccagaaccgt ttccatggag atgctcaggt ggctctgctc ggcctagctt 1200
cggtgatcaa gaacacaccg ctgtactttg ctgacaaact tcatcaagcc ctccaggaaa 1260
ctgagcccaa ttaccaagtc ctgattcgca tccttatctc tcgatgtgag actgaccttc 1320
tgagtatcag agctgagttc aggaagaaat ttgggaagtc cctctactct tctctccagg 1380
atgcagtgaa aggggattgc cagtcagccc tcctggcctt gtgcagggct gaagacatgt 1440
gagacttccc tgccccaccc cacatgacat ccgaggatct gagatttccg tgtttggctg 1500
aacctgggag accagctggg cctccaagta ggataacccc tcactgagca ccacattctc 1560
tagcttcttg ttgaggctgg aactgtttct ttaaaatccc ttaattttcc catctcaaaa 1620
ttatatctgt acctgggtca tccagctcct tcttgggtgt ggggaaatga gttttctttg 1680
atagtttctg cctcactcat ccctcctgta ccctggccag aacatctcac tgatactcga 1740
attcttttgg caaacttcgc tgttgtttgt gttccctgat tgaaggttgg gtggagcagg 1800
acatggaccg ggaagaggca ctggagttgg aggtgcct 1838

<210> 65
<211> 1689
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2658329CB1

<400> 65
tcacgtcggc cggagctagc gctgcgtcct gggccacgcc tcccggcgca ccgacgcgcc 60
tctccggtta ctaagcggcc ttggatacct ggccgcggga tgctgggcgg cgtcaggtaa 120
ccatggagaa agagctgcgg agcaccattc ttttcaatgc ctacaaaaag gagatattta 180
ccaccaacaa tggctacaaa tccatgcaga aaaaacttcg gagtaattgg aagattcaga 240
gcttaaaaga tgaaatcaca tctgagaagt taaatggagt aaaactgtgg attacagctg 300
ggccaaggga aaaatttact gcagctgagt ttgaaatcct gaagaaatat cttgacactg 360
gtggagatgt ctttgtgatg ctaggagaag gtggagaatc cagatttgac accaatatta 420
actttttact agaagaatat ggaatcatgg ttaataatga tgctgtggtt agaaatgtat 480
atcacaaata tttccatcct aaagaagctc tagtttccag tggagtcttg aacagggaaa 540
ttagccgagc tgcaggaaag gctgtgcctg ggatcattga tgaggaaagc agtggaaaca 600
atgcccaggc tctcaccttt gtgtatcctt ttggtgccac attgagtgtc atgaaaccag 660
cagtggcggt tctgtctaca ggttctgtct gcttcccact taacagaccc attttggctt 720
tctatcactc aaagaaccaa ggtgggaagc tggcagtgct tggttcatgt cacatgttca 780
gtgatcaata tttggacaaa gaagaaaaca gcaaaatcat ggatgttgtt ttccagtggc 840
tcacgacagg agacatccac ctaaaccaga ttgatgctga ggacccagag atttctgact 900
acatgatgct gccctacaca gccaccccta caaagcggaa tcgagagtgt ctccaggaga 960
gtgatgagat cccaagggac tttaccaccc tcttcgacct gtccatcttc cagctggata 1020
ccacctcctt ccacagcgtc atcgaggctc acgagcagct aaatgtgaaa catgaaccac 1080
tccagctcat ccagcctcag tttgagacgc cgctgccaac ccttcagcct gcggtttttc 1140
ctcccagttt ccgggagtta ccacctcctc ctctggagct atttgattta gatgaaacgt 1200
tctcctctga gaaggcacgg ctggctcaga ttaccaataa gtgtactgaa gaagacctgg 1260
aattttatgt caggaagtgt ggtgatattc ttggagtaac cagtaaacta ccaaaggacc 1320
aacaggatgc caaacatatc cttgagcacg tcttcttcca agtggtggag ttcaagaaat 1380
tgaaccagga acatgacatc gatacaagtg aaacagcatt ccagaacaat ttctgaagac 1440
catgcctctt gaagcttttt ctgcctcctg attctctctt tgtaaactat tttcaaattg 1500
tttttcaact ccttatcaaa attgtttata cactctttcc tccatgagct ctggaaggta 1560
tatgcatctt ctgtaatact cagataggta taagattttt cacaaaatcc ttatgtaaga 1620
tacattccat ttttaaaaat taaatgtatg gttgcatctg tctttttata ccctaaaaaa 1680
aaaaaaaaa 1689

<210> 66
<211> 1788
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 2708944CB1

<400> 66
cgagctcgcc cgctgtccgc cagcccgcgg gagggaggag agaagcgacg atgtccgcgg 60
ttggctactc agtgtcttgg tctcaagttg cctcattgcg gctggcgttc ccaatacaga 120
cgcatcgttt cttttttaat actccctaag aaagggaata accttcaagc tggcgggagc 180
aatggttcac ataaagaaag gcgagctgac ccaggaggag aaggagctac tggaagtcat 240
cgggaaaggt actgtccaag aagctggaac attattatcc agcaagaatg ttcgtgtcaa 300
ctgtttggac gagaatggaa tgactcctct aatgcatgca gcatataaag gaaaactgga 360
tatgtgcaaa ttactcctgc gacatggagc cgatgtaaat tgtcatcagc atgaacatgg 420
atacacagcc ctcatgtttg ctgcactttc tggtaataaa gacatcacat gggtaatgtt 480
agaagctggt gctgagacag atgttgtcaa ctctgtggga agaacagcag ctcagatggc 540
agcctttgtg ggtcaacatg attgtgtgac cataatcaac aatttctttc ctcgagagag 600
actggattat tacactaagc cccagggact ggataaagag ccaaaactgc cccccaaagtt 660
ggcaggcccg ctgcacaaaa ttatcaccac aacgaatctt catcctgtca agatcgtgat 720
gcttgtaaat gagaatcctc tgctgacaga agaagcagcc ctgaataaat gctacagagt 780
gatggatttg atttgtgaga aatgtatgaa gcaaagagac atgaatgaag tattggctat 840
gaagatgcat tacataagct gtatctttca gaaatgcatt aacttcttaa aagatggaga 900
gaataaactg gacaccttga tcaaaagctt gttaaaaggc cgagcttctg atggctttcc 960
agtgtatcaa gaaaagatca ttagagaaag tatcagaaaa tttccttact gtgaagctac 1020
actcctccag cagctggttc gaagcattgc tcctgttgaa attggttctg atcccactgc 1080
attctccgtc cttacccaag ccatcactgg ccaggtgggt tttgtggatg tggaattttg 1140
cactacctgt ggagaaaagg gagcaagtaa aagatgttca gtttgcaaaa tggtaatata 1200
ttgtgatcaa acctgccaga aaacacactg gtttactcat aagaaaatct gtaagaatct 1260
gaaggacatt tacgaaaagc aacagttgga ggctgccaaa gaaaagagac aagaggaaaa 1320
ccacggcaaa cttgatgtca attctaactg tgttaatgaa gagcaaccag aggctgaagt 1380
aggtatctct caaaaggatt ccaatcctga agattccggg gaaggaaaga aagaatctct 1440
tgaaagcgaa gctgagttgg aaggcttaca ggatgctcct gcagggccac aggtgtctga 1500
ggagtaaaag ccagagcaag tgccagtgtg gatggtcctc accctgcaag aagctggaaa 1560
actcctagga atgcattgtc ctcaccttgt tatacctgcg tggcaccatg gcaggattcc 1620
acatttcata gaatacaggt tttcaagcaa accctgttg accatgccct aatttcctat 1680
tgatttctgt tgtataattg aatggatatt cctatggaaa attttttgtt tcaaaataca 1740
ggaaaaacat tcctattacc tttctgaggc tggctttcca gcaattgt 1788

<210> 67
<211> 2160
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3315012CB1

<400> 67
atgctacgcg cgcccggctg cctcctccgg acctccgtag cgcctgccgc ggccctggct 60
gcggcgctgc tctcgtcgct tgcgcgctgc tctcttctag agccgaggga cccggtggcc 120
tcgtcgctca gcccctattt cggcaccaag actcgctacg aggatgtcaa ccccgtgcta 180
ttgtcgggcc ccgaggctcc gtggcgggac cctgagctgc tggaggggac ctgcaccccg 240
gtgcagctgg tcgccctcat tcgccacggc acccgctacc ccacggtcaa acagatccgc 300
aagctgaggc agctgcacgg gttgctgcag gcccgcgggt ccagggatgg cggggctagt 360
agtaccggca gccgcgacct gggtgcagcg ctggccgact ggcctttgtg gtacgcggac 420
tggatggacg ggcagctagt agagaaggga cggcaggata tgcgacagct ggcgctgcgt 480
ctggcctcgc tcttcccggt tcttttcagc cgtgagaact acggccgcct gcggctcatc 540
accagttcca agcaccgctg catggatagc agcgccgcct tcctgcaggg gctgtggcag 600
cactaccacc ctggcttgcc gccgccggac gtcgcagata tggagtttgg acctccaaca 660
gttaatgata aactaatgag atttttttgat cactgtgaga agtttttaac tgaagtagaa 720
aaaaatgcta cagctcttta tcacgtggaa gccttcaaaa ctggaccaga aatgcagaac 780
atttaaaaa aagttgcagc tactttgcaa gtgccagtaa atgatttaaa tgcagattta 840
attcaagtag cctttttcac ctgttcattt gacctggcaa ttaaaggtgt taaatctcct 900
tggtgtgatg tttttgacat agatgatgca aaggtattag aatatttaaa tgatctgaaa 960
caatattgga aaagaggata tgggtatact attaacagtc gatccagctg caccttgttt 1020
caggatatct ttcagcactt ggacaaagca gttgaacaga aacaaaggtc tcagccaatt 1080
tcttctccag tcatcctcca gtttggtcat gcagagactc ttcttccact gctttctctc 1140
atgggctact tcaaagacaa ggaacccta acagcgtaca attacaaaaa acaaatgcat 1200
cggaagttcc gaagtggtct cattgtacct tatgcctcga acctgatatt tgtgctttac 1260
cactgtgaaa atgctaagac tcctaaagaa caattccgag tgcagatgtt attaaatgaa 1320

aaggtgttac ctttggctta ctcacaagaa actgtttcat tttatgaaga tctgaagaac 1380
cactacaagg acatccttca gagttgtcaa accagtgaag aatgtgaatt agcaagggct 1440
aacagtacat ctgatgaact atgagtaact gaagaacatt tttaattctt taggaatctg 1500
caatgagtga ttacatgctt gtaataggta ggcaattcct tgattacagg aagcttttat 1560
attacttgag tatttctgtc ttttcacaga aaaacattgg gtttctctct gggtttggac 1620
atgaaatgta agaaaagatt tttcactgga gcagctctct taaggagaaa caaatctatt 1680
tagagaaaca gctggccctg caaatgttta cagaaatgaa attcttccta cttatataag 1740
aaatctcaca ctgagataga attgtgattt cataataaca cttgaaaagt gctggagtaa 1800
caaaatatct cagttggacc atccttaact tgattgaact gtctaggaac tttacagatt 1860
gttctgcagt tctctcttct tttcctcagg taggacagct ctagcatttt cttaatcagg 1920
aatattgtgg taagctggga gtatcactct ggaagaaagt aacatctcca gatgagaatt 1980
tgaaacaaga aacagagtgt tgtaaaagga caccttcact gaagcaagtc ggaaagtaca 2040
atgaaaataa atatttttgg tatttattta tgaaatattt gaacattttt tcaataattc 2100
ctttttactt ctaggaagtc tcaaaagacc atcttaaatt attatatgtt tgggacaata 2160

<210> 68
<211> 1156
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4155412CB1

<400> 68
ctgagacgag cgggagcgcg gacagcagcc tctgctgccc tgacttttta agaaatctca 60
atgaactatt tgtagagaat cactgatccg gcctgcaagc attttgcacg gcaaaaatat 120
cgatcagtgt taagtgaaga tcacatttta tatgcgatct tgactttttt gtcttacatt 180
atatttttat agattttgtt ataaacatgg tgctgggaaa ggtgaagagt ttgacaataa 240
gctttgactg tcttaatgac agcaatgtcc ctgtgtattc tagtggggat accgtctcag 300
gaagggtaaa tttagaagtt actggggaaa tcagagtaaa atctcttaaa attcatgcaa 360
gaggacatgc gaaagtacgc tggactgaat ctagaaacgc cggctccaat actgcctata 420
cacagaatta cactgaagaa gtagagtatt tcaaccataa agacatctta attgggcacg 480
aaagagatga tgataattcc gaagaaggct tccacactat tcattcagga aggcatgaat 540
atgcattcag cttcgagctt ccacagacac cactcgctac ctcattcgaa ggccgacatg 600
gcagtgtgcg ctattgggtg aaagccgaat tgcacaggcc ttggctacta ccagtaaaat 660
taaagaagga atttacagtc tttgagcata tagatatcaa cactccttca ttactgtcac 720
cccaagcagg cacaaaagaa aagacactct gttgctggtt ctgtacctca ggcccaatat 780
ccttaagtgc caaaattgaa aggaagggct ataccccagg tgaatcaatt cagatatttg 840
ctgagattga gaactgctct tcccgaatgg tggtgccaag gcagccattt accaaacaca 900
ggccttctat tgctaaaggg aaattgaggg agctaaacag cttgtggcta acatgcgtgg 960
ggaattcctt aacatctgga aagaaccggg acgtggaaat ggccagtttg ctgaaaattt 1020
ccaacagttt tccccctcc aatgcttcga actgaaggaa taatcccggc tgggaataat 1080
ccactcaatg ggtaaaatgg tgggaaaatc ccctggagcc aatggaattt aaattcctct 1140
aaatttggcc cacttg 1156

<210> 69
<211> 1981
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4831840CB1

<400> 69
ggctggggaa gatggcggtg gctggggcgg tgtccgggga gccgctggtg cactggtgca 60
cccagcagtt gcggaagact ttcggcctgg atgtcagcga ggagatcatt cagtacgttt 120
tgtcaattga gagtgctgaa gagatacgag aatatgttac tgatctcctc cagggaaatg 180
aaggcaaaaa aggtcaattc atagaagaac ttataaccaa atggcaaaag aatgatcagg 240
agttgatttc ggatcctttg cagcagtgct tcaaaaaaga tgaaatttta gatgggcaga 300
aatcaggcga ccatctaaag cggggtagga agaaagggag aaacagacag gaagttcctg 360
catttactga acctgacacg actgcagagg ttaaaacacc ttttgatttg gccaaggcac 420
aagagaacag caactccgta aagaagaaga caaagtttgt caatttatac acaagagagg 480
gacaggacag gcttgcagtc ctgctccctg gtcgtcaccc ttgtgattgc ctgggccaga 540
agcacaagct catcaataac tgtctgatct gtgggcgcat tgtctgtgaa caagaaggct 600
caggcccttg cttattctgt ggcactctgg tgtgtactca tgaggaacaa gatattttac 660

agcgtgactc aaacaagagc cagaaactgc taaagaaact catgtcagga gtggagaatt 720
ctggaaaggt ggacatctct accaaggacc ttcttcctca tcaagaattg cgaattaagt 780
ctggtctgga gaaggctatc aagcataaag acaaactgtt agagtttgac agaactagta 840
ttcgaaggac ccaagtcatt gatgatgagt cagattactt tgccagtgat tctaaccaat 900
ggttgtccaa acttgagcgg gaaaccttgc agaagcgaga ggaggagctg agagaacttc 960
gacacgcctc tcgactttct aagaaggtca ccattgactt tgcaggaagg aagatcctgg 1020
aagaagaaaa ttcactagca gagtatcata gcagactaga tgagacaata caggccattg 1080
ccaatggaac cttgaaccag ccactgacca aattggatag atcttctgaa gagcctttgg 1140
gagttctggt aaatcccaac atgtaccagt cccctcccca gtgggttgac cacacaggtg 1200
cagcctcaca gaagaaggct ttccgttctt caggatttgg actagagttc aactcatttc 1260
agcaccagtt gcgaatccag gatcaagaat ttcaggaagg ctttgatggt ggctggtgcc 1320
tctctgtaca tcagccctgg gcttctctgc ttgtcagagg gattaaaagg gtggagggca 1380
gatcctggta caccccccac agaggacgac tttggatagc agccacagct aaaaaaccct 1440
cccctcaaga agtctcagaa ctccaggcta catatcgtct tcttcgtggg aaagatgtgg 1500
aatttcctaa tgactatccg tcaggttgtc ttctgggctg tgtggaccta attgactgct 1560
tgtcccagaa gcaatttaag gagcagtttc cagacatcag tcaagaatct gattctccat 1620
ttgttttcat ctgcaaaaat cctcaggaaa tggttgtgaa gtttcctatt aaaggaaatc 1680
caaaaatctg gaaattggat tccaagatcc atcaaggagc aaagaagggg ttaatgaagc 1740
agaataaagc tgtctgaccc aggagaaaag gaactataca gcatagtgga gttttgtgta 1800
ctaaaattgc tatctactgg tcctttggaa ttgaagtagt agaaacctaa aggcttggcg 1860
tcaggcttga atatctcaga acttaaactc ttaccaaaat ctgtatattt ttcttaagga 1920
gtgggattcc tactttatgt aatggggtcg aaatctttga acacattatt tataaaaacc 1980
a 1981

<210> 70
<211> 1832
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5676581CB1

<400> 70
cctagagttg ttcagaatcc tccaaaacca gtcatgacca ctagacccac agctgttaaa 60
gcaacaggcg gtctatgctt gcttggtgct tatgctgaca gtgatgacga tgacaatgat 120
gtttccgaaa aactagcaca atccaaagag acaaatggaa accagtcaac tgatattgat 180
agtacattgg ccaacttcct agcggagatc gatgccataa cagctcctca gcctgcagct 240
cctgtaggag cttctgctcc acctccaact ccacctcgac cagagccaaa ggaagcagca 300
acatctaccc tttcttcttc tacttcaaat ggaacagact ccacccaaac atctggttgg 360
caatatgata ctcagtgttc actggcagga gtcggaattg agatgggcga ttggcaggaa 420
gtctgggatg agaacacggg atgttattat tattggaata cacaaacaaa tgaagtgact 480
tgggagttac cccaatatct tgccacacag gtacagggat tacagcatta ccagcccagt 540
tctgtgccag gtgctgaaac tagttttgtg gtaaatacag acatatattc taaggagaaa 600
acgatttctg tttccagtag taaaagtgga ccagtcatag ccaagcgaga agttaaaaag 660
gaagtaaatg aaggaattca ggctctctca aatagtgagg aggagaagaa aggggtggca 720
gcatcgctgc ttgctccttt attgcctgag ggaataaaag aagaagaaga gagatggaga 780
agaaaagtaa tttgtaaaga ggagccagtt tcagaagtaa aagaaacaag tacaacagta 840
gaagaagcaa caacaatagt aaagccacag gaaattatgt tggacaatat agaagaccct 900
tctcaggagg atctttgcag tgttgtccaa tctggagaaa gtgaggagga agaggaacaa 960
gataccottg aactggagct agttttggaa aggaaaaaag cagagttgcg agccttggag 1020
gaaggagatg gtagtgtgtc agggtctagt ccacgttctg atatcagcca gccagcatct 1080
caagatggaa tgcgtaggct tatgtctaaa agaggaaaat ggaagatgtt tgttcgagct 1140
accagtccag aatctaccag taggagttct agtaaaactg gacgagatac tccagaaaat 1200
ggagaaactg caattggtgc tgaaaattca gaaaaaatag atgagaattc agataaagag 1260
atggaagtag aagaatctcc agagaaaata aaagtacaga caacaccaaa agtagaagaa 1320
gaacaggatt tgaaatttca gattggagaa ctggcaaata ccctgacaag taaattcgag 1380
tttctaggca ttaatagaca atccatctcc aactttcatg tgctgctctt acagactgag 1440
actcgaattg cagactggcg ggaaggggct cttaatggaa actaccttaa acgaaaactt 1500
caggatgcag cagaacaact aaaacagtat gaaataaacg ccactcctaa aggctggtcc 1560
tgccactggg acaggtacgc actcttctcc ccttttcacc tttcaccttt gacatctcag 1620
acatgatttg tgatcaccac catctgacga catgacagcc tgtctggaga ctgagagcag 1680
ctgcagttag cggtgctggg caggcaaagc accgcaagca caaagtttgg ccagccacac 1740
attggagaaa caaaagcaat gtttaaagtg ttccatgtaa cggatttttt cccaagatat 1800
ggacaaagct ggtttttact ctccagagtg tt 1832

<210> 71

<211> 1772
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 034159CB1

<400> 71
acagtactga tattaagcag catccaacac aggcctactc ttacgacatg tgactttact 60
gttttccgtt tttgttgaaa gagtcattaa cagttaggag ttgatggcag tttcaataac 120
aggtcattgc cgagaaaagg atagcactat aatatgcaga aatctacaaa ttctgatact 180
tccgtggaaa cactgaattc tacccgccaa ggcacaggag ctgtgcaaat gagaatcaaa 240
aatgccaaca gccaccatga caggctcagc caaagtaaat ccatgatcct caccgatgtc 300
gggaaggtca ctgaacctat atccagacac agaaggaatc attcacagca tatcttgaaa 360
gatgtcattc ctccattgga acaactgatg gttgaaaaag aaggttatct gcaaaaagct 420
aaaattgcag atggaggaaa gaaactaagg aaaaactggt ctacttcctg gattgttctt 480
tctagtcgaa gaattgaatt ttacaaagaa tccaagcaac aggctctgtc caatatgaaa 540
actgggcaca aaccagaaag tgtggatttg tgtggagcac acattgaatg ggccaaggaa 600
aaatcgagca gaaagaatgt ctttcagatc acaacagtat caggaaatga gttccttcta 660
cagtcagata ttgacttcat catattggat tggttccacg ctatcaaaaa tgcaattgac 720
agattgccaa aggattcaag ttgtccatca agaaacctgg aattattcaa aatccaaaga 780
tcctctagca ctgaattgct aagtcactat gacagtgata taaaagaaca gaaaccagag 840
cacagaaaat ctttaatgtt cagactgcat cacagtgctt ccgatacaag cgacaaaaat 900
cgagttaaaa gcagattaaa gaagtttatt acccgaagac cttccctgaa aactctgcaa 960
gaaaaaggac ttattaaaga tcaaattttt ggctctcatc tgcacaaagt gtgtgaacgt 1020
gaaaattcca cagttccgtg gtttgtaaag caatgcattg aagctgttga gaaaagaggt 1080
ctagatgttg atggaatata tcgagttagt ggcaatctgg caacaataca gaagttaaga 1140
tttattgtca accaagaaga gaagctgaat ttggacgaca gccagtggga ggacatccac 1200
gttgtcaccg gagcactgaa gatgtttttc cgggagctgc ctgagccgct cttcccttac 1260
agtttctttg agcagtttgt ggaagcgatc aaaaagcaag acaacaacac aagaattgaa 1320
gctgtaaaat ctcttgtaca aaaactccct ccgccaaatc gtgacaccat gaaagtcctc 1380
tttggacatc taactaagat agtggccaaa gcctccaaga acctcatgtc cacgcaaagc 1440
ttggggattg tatttggacc taccctctg cgagctgaaa atgaaacagg aaacatggcg 1500
atccacatgg tctaccagaa ccagatagct gagctcatgc tgagtgagta cagtaagatc 1560
ttcggctcag aggaagactg acagacaaga caagctactg aatacgttca catctgtctt 1620
gatgcctaat atttttacat ttctgtaaac atatttctga aatatttttt gcctttcaag 1680
cgacagatgc ctcattttgt gaaaacttaa tgatgatttt gtgtttaagt tccaaacatt 1740
tgaataaaat aattgacaat aaaaaaaaaa aa 1772

<210> 72
<211> 1488
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 129023CB1

<400> 72
cgggacacaa gatggcggca gcggcgctgg ggagggcgag gcggaggcgg caaaacgggc 60
ggtcgagcag aacgtgtagc cgcgtcccct ccagtccgct ccgggcagct gctgatgcaa 120
ggaatcccct gggctcccgt ccactccact gctgaccagc ccattcgcct gtgctgagtc 180
ttcctgcagg cctttccttg cctctgtggg accctgtggg ggtccatccg gctggagaag 240
aaaagcctct catgctaacg ttgcagaccc cagagggtcc tgtgtgggtg tggagatggc 300
caatgagaat cacggcagcc cccgggagga agcgtccctg ctgagtcact ccccaggtac 360
ctccaatcag agccagccct gttctccaaa gccaatccgc ctggttcagg acctcccaga 420
ggagctggtg catgcaggct gggagaagtg ctggagccgg agggagaatc gtccctacta 480
cttcaaccga ttcaccaacc agtccctgtg ggagatgccc gtgctggggc agcacgatgt 540
gatttcggac cctttggggc tgaatgcgac cccactgccc caagactcaa gcttggtgga 600
aactcccccg gctgagaaca agcccagaaa gcggcagctc tcggaagagc agccaagcgg 660
caatggtgtg aagaagccca agattgaaat cccagtgaca cccacaggcc agtcggtgcc 720
cagctccccc agtatcccag gaaccccaac gctgaagatg tggggtacgt cccctgaaga 780
taaacagcag gcagctctcc tacgacccac tgaggtctac tgggacctgg acatccagac 840
caatgctgtc atcaagcacc gggggccttc agaggtgctg cccccgcatc ccgaagtgga 900
actgctccgc tctcagctca tcctgaagct tcggcagcac tatcgggagc tgtgccagca 960
gcgagagggc attgagcctc cacgggagtc tttcaaccgc tggatgctgg agcgcaaggt 1020

ggtagacaaa ggatctgacc ccctgttgcc cagcaactgt gaaccagtcg tgtcaccttc 1080
catgtttcgt gaaatcatga acgacattcc tatcaggtta tcccgaatca agttccggga 1140
ggaagccaag cgcctgctct ttaaatatgc ggaggccgcc aggcggctca tcgagtccag 1200
gagtgcatcc cctgacagta ggaaggtggt caaatggaat gtggaagaca cccttagctg 1260
gcttcgcgaa ggaccactca gcctccaaag aggggcaaca tggatcgcct ggaacatctg 1320
cggaagcagt gctggcccca agtctcggcc gcaaccaaga ctccgtgcga acgcatctgc 1380
agtagatcta caaatctcct gaggtagtaa acgatccgga gaagacctgc atcctcagaa 1440
acactctcca gagtgaggcc ctagtgagcc gctatgttgt accatcgc 1488

<210> 73
<211> 2430
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1358940CB1

<400> 73
ggcccagcgg ctaggagagt cacgtgagat tgggcggagg gggtggaggt ttgtctccgc 60
tgtttcatct ctatggctgt cagaggtggg cggctttgac cgagaggctg ctggagctcg 120
tgtttggacg cgatgtttcg tctgaactca ctttctgctt tggcagaact ggctgtgggt 180
tctcgatggt accatggagg atcacagccc atccagatcc ggcgaagact aatgatggtg 240
gctttcctgg gagcatctgc agtaactgca agtactggtc ttttgtggaa gagggcccat 300
gcagaatctc caccatgtgt agacaaccta aaaagtgaca tcggtgataa agggaagaat 360
aaagatgaag gggatgtttg taaccatgag aaaaagactg cagatcttgc ccctcaccca 420
gaagagaaaa agaagaaacg ttctggattc agagacagaa aagtgatgga atatgagaat 480
aggattcgag cctactccac gccagacaaa atcttccgat attttgccac cttgaaagtc 540
atcagtgagc ctggtgaagc agaagtgttt atgacaccag aagattttgt gcgatccata 600
acacccaatg aaaaacaacc agaacacttg ggtctggatc aatatataat aaaacgcttt 660
gatggaaaga aaatttccca ggaacgagaa aaatttgctg atgaaggcag tatattttac 720
acccttggag aatgtgggct catatccttt tcagactaca ttttcctcac aactgttctt 780
tccactcctc agagaaattt tgaaattgcc ttcaagatgt ttgatttgaa tggagatgga 840
gaagtagata tggaagaatt tgaacaggtt cagagcatca ttcgctccca aaccagtatg 900
ggtatgcgcc acagagatcg tccaactact ggcaacaccc tcaagtctgg cttgtgttca 960
gccctcacaa cctacttttt tggagctgat ctgaagggaa agctgacaat caaaaacttc 1020
ctcgaatttc agcgtaaact gcagcatgat gttctgaagc ttgagtttga acgccatgac 1080
cctgtggatg ggagaattac tgagaggcag tttggtggca tgctacttgc ctacagtggg 1140
gtgcagtcca agaagctgac cgccatgcag aggcagctca agaagcactt caaagaagga 1200
aagggtctga catttcagga ggtggagaac ttctttactt tcctaaagaa cattaatgat 1260
gtggacactg cattgagttt ttaccatatg gctggagcat ctcttgataa agtgaccatg 1320
cagcaggtgg ccaggacagt ggctaaagtg gagctctcag accacgtgtg tgatgtggtg 1380
tttgcactct ttgactgtga tggcaatggc gaactgagca ataaggaatt tgtttccatc 1440
atgaagcaac ggctgatgag aggcctggaa aagcccaaag acatgggttt cactcgcctc 1500
atgcaggcca tgtggaaatg tgcacaggaa actgcctggg acttcgcttt acccaaacag 1560
taaccccaca ctgcaagagg ggacccctcc acccccagta ccctggaccc cctccgcaga 1620
gtctcggcag agccctttgt gctgctgctt ctggaagtag tctcccttcc tcccgggatg 1680
acctcaggac tctgtcggtt tcccctcttt acccttcccc gtccccgtgt tctgctgggc 1740
tctgattctg cccaatgagt atccccatag gttctcaaaa acatgaacaa gtctgtaaag 1800
ctcagacatt tgtcagcctc aacagcacca cccattcaag catcctgtgg ataaagaatt 1860
cagggaacca tccacacacc tgccaaccct gggaagcatc cagttctcaa atcgtttttg 1920
ctatggattt atactaacaa gaacattcct tgacttccct cctgctggtg ttttaaagcc 1980
acaagtaggg aagatatctg gcaggcagaa agaagtctgt gatgataaac aatgatgagg 2040
atgacctagg caccctacgc tagtgtgaga agcctgcgcc ccaggaagga tctgtgttag 2100
tccctgggat ggctccaagg cctgctctag gaaggcagca tgctcagtgg gaacacagca 2160
agattcagaa tttaaagtag ttgcttcatg gctctgtgca ctccctttc ttcctcgcag 2220
cctccctaag atgactccag tgtgaccctg tgcttagtga gcaatagtga ttgagctcat 2280
gttccctgca agtgccattt cctctccagg atgggcctct aaagctgagg cctggctcag 2340
agcctgtttg ccctctgtct taaacaattg taaatatcac ttaaattata accatttgca 2400
ataaacatcc ccaaagttaa aaaaaaaaaa 2430

<210> 74
<211> 1411
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 1682320CB1

<400> 74
agcactagtc gcgcccactc ccttctactt ccaggtcggg ggggggcggg tccaatagaa 60
aggcggaagc cagtgtccca ggcgttctca cgcccgcaac aattcctgag tagggccttg 120
cttgagttct tcggaaagtc tcatccaccc ccacatcgcc tctttaggaa gtcacttaat 180
gttgggcttc attattccca catccctttc cttactactt gcctgcactt cttgagaaaa 240
agactgcaga aaggagaggt ggggctttca gtagaaacaa gcaaaccgca ggtccctgtg 300
gggggactct ccaggaagaa gggtaatttc ctgcctcctt aaattggctg ctactgtcag 360
ttattttgct cccaacccca gagcttcact tgctccttca cttcccagtt ccgcaagaac 420
cgtgggcgac agttatggag aagcgtctgc aggaggctca gctgtacaag gaggaaggga 480
accagcgcta ccgggaaggg aagtaccgag atgctgtgag taggtaccat cgagctctgc 540
ttcagctgcg gggtctggat ccgagtctgc cctctccgtt acctaatctc ggacctcagg 600
gcccggccct cacgcctgaa caagaaaaca tattgcatac cacccagaca gactgctata 660
acaatctagc tgcttgtctc cttcagatgg agcccgtgaa ctacgaacga gtgagagaat 720
atagtcagaa agtcctggaa cgacagcctg ataatgccaa ggccttgtat cgggccggag 780
tggccttttt ccatctgcag gactatgacc aggcccgcca ctacctcctg gctgccgtga 840
ataggcagcc taaagatgcc aacgtccggc ggtacctcca gctgacacag tcagaactca 900
gcagctacca tagaaaagag aagcagctct acctgggcat gtttggttaa caaagaagaa 960
agatgctcct ccagttgaac ttaggtggac cattaaacat gcatgaagga gaaatctgag 1020
cctcagcaag agaaattaac cctatacctc tgacccaggt ggatttttgt ttctagttct 1080
gcacaaactt cactacttag acagtctgag tctttttctg tctatccatc tgtttatttc 1140
tatacctttc aatacatgtt attgttgcag atatttggct tgagaaatat aatcagaaaa 1200
catacatcag ttgtgggtgg aattaatcat atctctggta tagatttttc atgacagtgt 1260
tgttagatgt acttatatac agaggcgaca gcttgcagag gacgacatag taaggatagg 1320
cagaaagaat tgtcttcctt tatttttca gaagcccaaa ccaaagtggg aaatatggcc 1380
gggcgctcta gctcatgcct gtaatcccag c 1411

<210> 75
<211> 653
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1728263CB1

<400> 75
gcggtggtag ctggttgggt tgctggggct cggttgttgt agtcgcgatg ttcttctccg 60
aggccagagc caggtcgcgg acgtgggaag ccagtccctc ggaacacagg aagtgggtgg 120
aagtatttaa agcatgtgat gaagatcaca aaggatatct cagcagagag gactttaaaa 180
ctgctgttgt aatgctgttt gggtacaagc cctccaagat agaagtggat tctgtgatgt 240
cttcaataaa tccaaatact tctggtatat tactcgaggg gtttttaaat attgtcagga 300
aaaagaagga agctcaacga tatcggaacg aagtaagaca catcttcaca gcctttgaca 360
cctactatcg tggattttta actttggaag atttcaaaaa agcatttagg caggtggctc 420
ccaaattacc ggaaaggact gttcttgagg tattcaggga agtagatcga gattcagatg 480
gtcacgtcag ctttagagac tttgaatatg ccctgaacta tggacagaag gaagcctaac 540
tattgtgaac tacttttggt aactctgggg agatcaatag attgtaatgt cagcagactc 600
tactctacta atgatgtcat gctacagact tgtgattaaa catttaaaaa ttt 653

<210> 76
<211> 1448
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1867626CB1

<400> 76
aacccggatt ctgcatttcc cctaactctg agactcattt tgtggaatag agttgatcgc 60
tgtctcctcc gcaaagcatt ttaactcgaa taagcaaatg ccgcctctgt ttgaacgttt 120
tggtatttac aagagagaaa tcattttacc taagagaact aattgaattg gcagcatcct 180
tgaaatacct ccggacaagg atctgggggt gggggtggaa aagcaactgc gaaatagcag 240
acggagaaat tcctttggaa gttattccgt agcataagag ctgaaacttc agagcaagtt 300
ttcattgggc aaaatggggg aacaacctat cttcagcact cgagctcatg tcttccaaat 360

tgacccaaac acaaagaaga actgggtacc caccagcaag catgcagtta ctgtgtctta 420
tttctatgac agcacaagaa atgtgtatag gataatcagt ttagatggct caaaggcaat 480
aataaatagt accatcaccc caaacatgac atttactaaa acatctcaga ggtttggcca 540
gtgggctgat agccgggcaa acaccgttta tggattggga ttctcctctg agcatcatct 600
ttcgaaattt gcagaaaagt ttcaggaatt taaagaagct gctcgactag caaaggaaaa 660
atcacaagag aagatggaac ttaccagtac accttcacag gaatccgcag gcggggatct 720
tcagtctcct ttaacaccgg aaagtatcaa cgggacagat gacgaaagaa cacctgatgt 780
gacacagaac tcagagccaa gggctgaacc aactcagaat gcattgccat tttcacatag 840
ttcagcaatc agcaaacatt gggaggctga actggctacc ctcaaaggaa ataatgccaa 900
actcactgca gccctgctgg agtccactgc caatgtgaaa caatggaaac agcaacttgc 960
tgcctatcaa gaggaagcag aacgtctgca caagcgggtg actgaacttg aatgtgttag 1020
tagccaagca aatgcagtac atactcataa gacagaatta aatcagacaa tacaagaact 1080
ggaagagaca ctgaaactga aggaagagga aatagaaagg ttaaaacaag aaattgataa 1140
tgccagagaa ctacaagaac agagggattc tttgactcag aaactacagg aagtagaaat 1200
tcggaacaaa gacctggagg gacaactgtc tgacttagag caacgtctgg agaaaagtca 1260
gaatgaacaa gaagcttttc gcaataacct gaagacactc ttagaaattc tggatggaaa 1320
gatatttgaa ctaacagaat tacgagataa cttggccaag ctactagaat gcagctaagg 1380
aaagtgaaat ttcagtgcca attaattaaa agatacactg tctctcttca taggaatgtt 1440
tagctctg 1448

<210> 77
<211> 1538
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1990126CB1

<400> 77
gcgccccagc ccggccgcgg cacctccgcc tcgccgccgc taggtcggcc ggctccgccc 60
ggctgccgcc taggatgaat atcatggact tcaacgtgaa gaagctggcg gccgacgcag 120
gcaccttcct cagtcgcgcc gtgcagttca cagaagaaaa gcttggccag gctgagaaga 180
cagaattgga tgctcactta gagaacctcc ttagcaaagc tgaatgtacc aaaatatgga 240
cagaaaaaat aatgaaacaa actgaagtgt tattgcagcc aaatccaaat gccaggatag 300
aagaatttgt ttatgagaaa ctggatagaa aagctccaag tcgtataaac aacccagaac 360
ttttgggaca atatatgatt gatgcaggga ctgagtttgg cccaggaaca gcttatggta 420
atgcccttat taaatgtgga gaaacccaaa aaagaattgg aacagcagac agagaactga 480
ttcaaacgtc agccttaaat tttcttactc ctttaagaaa ctttatagaa ggagattaca 540
aaacaattgc taaagaaagg aaactattgc aaaataagag actggatttg gatgctgcaa 600
aaacgagact aaaaaaggca aaagctgcag aaactagaaa ttcatctgaa caggaattaa 660
gaataactca aagtgaattt gatcgtcaag cagagattac cagacttctg ctagagggaa 720
tcagcagtac acatgcccat caccttcgct gtctgaatga ctttgtagaa gcccagatga 780
cttactatgc acagtgttac cagtatatgt tggacctcca gaaacaactg ggaagttttc 840
catccaatta tcttagtaac aacaatcaga cttctgtgac acctgtacca tcagttttac 900
caaatgcgat tggttcttct gccatggctt caacaagtgg cctagtaatc acctctcctt 960
ccaacctcag tgaccttaag gagtgtagtg gcagcagaaa ggccagggtt ctctatgatt 1020
atgatgcagc aaacagtact gaattatcac ttctggcaga tgaggtgatc actgtgttca 1080
gtgttgttgg aatggattca gactggctaa tgggggaaag gggaaaccag aagggcaagg 1140
tgccaattac ctacttagaa ctgctcaatt aagtaggtgg actatggaaa ggttgcccat 1200
catgactttg tatttatata caattaactc taaataaagc aggttaagta tcttccatgt 1260
taatgtgtta agagactgaa aataccagcc atcagaaact ggcctttctg ccaataaagt 1320
tgcatggtaa atatttcatt acagaattta tgttagagct ttcatgccaa gaatgttttc 1380
ttacaaaatt ctctttttat tgaggtttca ctaataagca gcttctactt ttgagcctca 1440
acttaaagca gaactgtttt ctactggatt tttcattaac agcaagcttt ttcttttatg 1500
taaaataaat ctattgtgaa ttgaaaaaaa aaaaaaaa 1538

<210> 78
<211> 998
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2104180CB1

<400> 78

ggggcggtgc cgggggcggg gcgcggcggc tgtcagctga ctgtggcggc ggcggcctcg 60
aggtgacaac tgtctccgtc gcaggctccg gcgggggcgc aggaggtcgc ccggcgcgtc 120
actgtcgggt cggcgagcca cgggggccgc cgcagcacca tggcgaccac cgtcagcact 180
cagcgcgggc cggtgtacat cggtgagctc ccgcaggact tcctccgcat cacgcccaca 240
cagcagcagc ggcaggtcca gctggacgcc caggcggccc agcagctgca gtacggaggc 300
gcagtgggca ccgtgggccg actgaacatc acggtggtac aggcaaagtt ggccaagaat 360
tacggcatga cccgcatgga cccctactgc cgactgcgcc tgggctacgc ggtgtacgag 420
acgcccacgg cacacaatgg cgccaagaat ccccgctgga ataaggtcat ccactgcacg 480
gtgcccccag gcgtggactc tttctatctc gagatcttcg atgagagagc cttctccatg 540
gacgaccgca ttgcctggac ccacatcacc atcccggagt ccctgaggca gggcaaggtg 600
gaggacaagt ggtacagcct gagcgggagg caggggggacg acaaggaggg catgatcaac 660
ctcgtcatgt cctacgcgct gcttccagct gccatggtga tgccaccccca gcccgtggtc 720
ctgatgccaa cagtgtacca gcagggcgtt ggctatgtgc ccatcacagg gatgcccgct 780
gtctgtagcc ccggcatggt gcccgtggcc ctgcccccgg ccgccgtgaa cgcccagccc 840
cgctgtagcg aggaggacct gaaagccatc caggacatgt tccccaacat ggaccaggag 900
gtgatccgct ccgtgctgga agcccagcga gggaacaagg atgccgccat caactccctg 960
ctgcagatgg gggaggagcc atagagcctc tgcctcga 998

<210> 79
<211> 1086
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2122241CB1

<400> 79
gacttgctga ggaaaggcat acgttaatgt atgtgtgagg ttaacccagg aggttgaatt 60
ccaaaatttc cccttcccct ttcttccttt tactcaggtt ctgtttcaag ctgattccag 120
ccatttggga ttgccgccca ctgacaactt agtagggctt tcattataag caggcttgag 180
atatctgacc tttctctcga cctaaacacc tggggctgga aacatggcgg gagagattta 240
cagattatac ctggggcaac gaaccagaac gacattagag gagaaaagag gcctgtcgtt 300
tttattggct tctgcacgtc ttcctcccca gcttcggcca ctctcccctt cgcaccctct 360
cacctgacag aagggactat tcctagttta tgaggtggtt aaggatatcg gtggggtggg 420
ctggagcggt gtcgggttag gtctgagaga aggcctcgca caaaacactg tacaaacccg 480
aaaggaagtc tgagagacga accgccttcc tccctgaagc ttctagaact ggagcagaaa 540
gaaggtgtgg cccagggcca gccccgcctc ctccccgggc ggaagctgtg tcagttgccg 600
gaagtcggcg tgaggtgggg cttatgcggc ggcgtggtga aatagatatg gcgaccgagg 660
gggatgtgga gctggagttg gagactgaga ccagtggacc agagcggcct ccggagaagc 720
cacggaaaca tgacagcggt gcggcggact tggagcgggt caccgactat gcagaggaga 780
aggagatcca gagttccaat ctggagacgg ccatgtctgt gattggagac agaaggtccc 840
gggagcagaa agccaaacag gagcgggaga aagaactggc aaaagtcact atcaagaagg 900
aagatctgga gctaataatg actgagatgg agatatctcg agcagcagca gaacgcagtt 960
tgcgggaaca catgggcaac gtggtagagg cgcttattgc cctaaccaac tgatgcgtgc 1020
tttctcaaat atacctactg gattaattta tggcaataaa attttttttt gtcttttaaa 1080
aaaaaa 1086

<210> 80
<211> 2323
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2580428CB1

<400> 80
tacgccaccg cgtccgagcc caaggcaaga gcgctacgcg ccctgcccga agtgcaactg 60
tcagtgagcc tgcgcaggag gccaataggc tgccaatact ccttggactc cccgccaggt 120
ccctgctgtc agtgcgcctg cgcgcggtc cggcgccgag gttcttgact gctgtgccgg 180
acgccaggtg tagccatgca gcgagccgat tccgagcagc cctccaagcg tccccgttgc 240
gatgacagcc cgagaacccc ctcaaacacc ccttccgcag aggcagactg gtccccgggc 300
ctggaactcc atcccgacta caagacatgg ggtccggagc aggtgtgctc cttcctcagg 360
cgcggtggct ttgaagagcc ggtgctgctg aagaacatcc gagaaaatga aatcacaggc 420
gcattactgc cttgtcttga tgagtctcgt tttgaaaatc ttggagtaag ttccttgggg 480
gagaggaaga agctgcttag ttatatccag cgattggttc aaatccacgt tgatacaatg 540

```
aaggtaatta atgatcctat ccatggccac attgagctcc accctctcct cgtccgaatc 600
attgatacac ctcaatttca acgtcttcga tacatcaaac agctgggagg tggttactat 660
gtttttccag gagcttcaca caatcgattt gagcatagtc taggggtggg gtatctagca 720
ggatgtctag ttcacgcact gggtgaaaaa caaccagagc tgcagataag tgaacgagat 780
gttctctgtg ttcagattgc tggactttgt catgatctcg gtcatgggcc attttctcac 840
atgtttgatg gacgatttat tccacttgct cgcccggagg tgaaatggac gcatgaacaa 900
ggctcagtta tgatgtttga gcaccttatt aattctaatg gaattaagcc tgtcatggaa 960
caatatggtc tcatccctga agaagatatt tgctttataa aggaacaaat tgtaggacca 1020
cttgaatcac ctgtcgaaga ttcattgtgg ccatataaag ggcgtcctga aaacaaaagc 1080
ttcctttatg agatagtatc taataaaaga aatggcattg atgtggacaa atgggattat 1140
tttgccaggg actgccatca tcttggaatc caaaataatt ttgattacaa gcgctttatt 1200
aagtttgccc gtgtctgtga agtagacaat gagttgcgta tttgtgctag agataaggaa 1260
gttggaaatc tgtatgacat gttccacact cgcaactctt tacaccgtag agcttatcaa 1320
cacaaagttg gcaacattat tgatacaatg attacagatg ctttcctcaa agcagatgac 1380
tacatagaga ttacaggtgc tggaggaaaa aagtatcgca tttctacagc aattgacgac 1440
atggaagcct atactaagct gacagataac atttttctgg agattttata ctctactgat 1500
cccaaattga aagacgcacg agagatttta aaacaaattg aataccgtaa tctattcaag 1560
tatgtgggtg agacgcagcc aacaggacaa ataaagatta aaagggagga ctatgaatct 1620
cttccaaaag aggttgccag tgctaaaccc aaagtattgc tagacgtgaa actgaaggct 1680
gaagatttta tagtggatgt tatcaacatg gattatggaa tgcaagaaaa gaatccaatt 1740
gatcatgtta gcttctattg taagactgcc cccaacagag caatcaggat tactaaaaac 1800
caggtttcac aacttctgcc agagaaattt gcagagcagc tgattcgagt atattgtaag 1860
aaggtggaca gaaagagttt gtatgccgca agacaatatt ttgttcagtg gtgtgcagac 1920
agaaatttca ccaagccgca ggatggcgat gttatagccc cactcataac acctcaaaaa 1980
aaggaatgga acgacagtac ttcagtccaa aatccaactc gcctccgaga agcatccaaa 2040
agcagagtcc agctttttaa agatgaccca atgtgaatgt ctgtagtcag ttgtttacaa 2100
actccctctc ctgcacaatt catttagagg cttcaatcat agaattctgc aaattaatga 2160
caactcatgc tttaattttg tattttgaat gtacacgcat gctgaagcta agtaactttt 2220
aatcaaagaa ataagatggt attaggcaaa tcttactata ctatgaaaag cattaccttg 2280
cctattttta atattattaa agcctttctc cttcaaaaaa aaa                  2323
```

<210> 81
<211> 669
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3397189CB1

<400> 81

```
cccacgcgtc cgaacgccat ggctcccaag aagctgtcct gccttcgttc cctgctgctg 60
ccgctcagcc tgacgctact gctgccccag gcagacactc ggtcgttcgt agtggatagg 120
ggtcatgacc ggtttctcct agacggggcc ccgttccgct atgtgtctgg cagcctgcac 180
tactttcggg taccgcgggt gctttgggcc gaccggcttt tgaagatgcg atggagcggc 240
ctcaacgcca tacagtttta tgtgccctgg aactaccacg agccacagcc tggggtctat 300
aactttaatg gcagccggga cctcattgcc tttctgaatg aggcagctct agcgaacctg 360
ttggtcatac tgagaccagg accttacatc tgtgcagagt gggagatggg gggtctccca 420
tcctggttgc ttcgaaaacc tgaaattcat ctaagaacct cagatccagg tgagttgaga 480
caaaggattt aacacagaag caagtaagta aaatgggcta tttgggtgcc aaaagcagaa 540
gagaccattc ccaaattgga ggtcatcatt catttaccaa gtgtttcctt catgcccagc 600
aggatgctag aaactggggg accagacaga cccccatcct tgtccagcag gcttatgata 660
tggtcaatc                                                        669
```

<210> 82
<211> 1606
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4881249CB1

<400> 82

```
gcggcggaga cggccggcgc ccggcgcgga gcctagggag gcagttcagc gcggcctcgg 60
gcctcgtcga gaaggatgct gtcccgaaag aaaaccaaaa acgaagtgtc caagccggcc 120
gaggtgcagg ggaagtacgt gaagaaggag acgtcgcctc tgcttcggaa tcttatgcct 180
```

tcattcatcc ggcatggtcc aacaattcca agacgaactg atatctgtct tccagattca 240
agccctaatg ccttttcaac ttctggagat gtagtttcaa gaaaccagag tttccttaga 300
actccaattc aaagaacacc tcatgaaata atgagaagag aaagcaacag attatctgca 360
ccttcttatc ttgccagaag tctagcagat gtccctagag agtatggttc ttctcagtca 420
tttgtaacgg aagttagttt tgctgttgaa aatggagact ctggttcccg atattattat 480
tcagacaatt tttttgatgg tcagagaaag cggccacttg gagatcgtgc acatgaagac 540
tacagatatt atgaatacaa ccatgatctc ttccaaagaa tgccacagaa tcaggggagg 600
catgcttcag gtattgggag agttgctgct acatctttag gaaatttgac taaccatggt 660
tctgaagatt tacccccttcc tcctggctgg tctgtggact ggacaatgag agggagaaaa 720
tattatatag atcataacac aaatacaact cactggagcc atcctcttga gcgagaagga 780
cttcctcctg gatgggaacg agttgagtca tccgaatttg gaacctatta tgtagatcac 840
acaaataaga aggcccaata caggcatccc tgtgctccta gtgtacctcg gtatgatcaa 900
ccacctcctg tcacatacca gccacagcaa actgaaagaa atcagtccct tctggtacct 960
gcaaatccat atcatactgc agaaattcct gactggcttc aggtttacgc acgagcccct 1020
gtgaaatatg accacattct gaagtgggaa ctcttccagc tggctgacct ggatacatac 1080
cagggaatgc taaagttgct cttcatgaaa gaattggagc agattgttaa aatgtatgaa 1140
gcatacagac aagcccttct tacagagttg gaaaaccgaa agcagagaca gcagtggtat 1200
gcccaacaac atggaaaaaa tttttgagct gattttttaa aaatttaagt tttgtaagag 1260
ctttaaaata ttttcacaga taaaaaattg caaacaagta ctctggttaa taaatgctgc 1320
ttcctttgtg gaaattataa aattctaact ttacatgtat tttgttatta gaaattttct 1380
tttattgaat gagaaaaatt agtctatcat tttaagagcc aatatggcaa acactttcaa 1440
atactgtata ttaggaaact gttttggtat tcttgatgga aaaaaatgca gcggaaatgt 1500
cattatgaac agatgttaaa taggaaatta ttacttgtta acttcttaca gcagtagtac 1560
cttctttaaa aaaaaaaaga atctgcggta ttttttttaa aaaaaa 1606

<210> 83
<211> 1980
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 431871CB1

<400> 83
ggctgagcgt gtttacatcc gccgggtgcg cggcttcgcc gcccgaggtc gttcggctcg 60
ggtaccatcc tccgcgccat ggacaccagc gacctgttcg ccagctgcag gaagggggat 120
gtgggccgag tgcggtacct gctggagcag cgagacgtgg aggtgaatgt gcgggacaag 180
tgggacagca ccccccttgta ctatgcctgc ttgtgtgggc acgaggagct ggtactctac 240
cttctggcca atggagcccg ctgcgaggcc aacacccttcg atggtgagcg ctgcctctat 300
ggggcactga gtgaccccat ccgccgggct ctacgcgatt acaagcaggt cacggcttcc 360
tgcaggaggc gggattacta tgacgacttc ttgcagcggc ttctagagca gggcatccac 420
agtgacgtgg tctttgtagt acacgggaag ccattccggg tgcatcgctg cgtcctgggt 480
gcacgtagtg cctactttgc caacatgctg gacaccaaat ggaagggcaa gagtgtcgtg 540
gttctcaggc acccactgat caaccccgtg gcctttgggg ccctgctgca gtacctgtac 600
acaggccgcc tggacattgg cgtagagcat gtgagtgact gtgagcgcct ggccaagcaa 660
tgccagctgt gggacctgct cagcgacctg gaggccaagt gcgagaaggt gtctgagttt 720
gtggcgtcta agccaggcac gtgtgtgaag gtgctgacca tcgagccccc acctgcagac 780
ccccgcctcc gggaggacat ggcgctgctg gccgattgtg ccctgccccc cgagctccga 840
ggtgatcttt gggagctgcc cttcccttgt cctgacggct tcaacagctg ccctgacatc 900
tgcttccgag tggctggctg cagcttcctc tgccacaagg ccttttttctg tggccgcagt 960
gactacttcc gagccctgct ggatgaccac ttccgagaga gcgaggagcc agcgacctca 1020
gggggccccc cagccgtcac cctgcatggc atctcacccg acgtcttcac tcacgtgctc 1080
tactacatgt acagcgacca cactgagctg tcccccgagg cagcctatga tgtgctgagc 1140
gtcgccgaca tgtacctgct gccaggcctg aagaggctgt gcggccgcag cctggctcag 1200
atgctagacg aggacactgt ggtgggtgtg tggcgcgtgg ccaagctctt ccgcctggcg 1260
cggcttgagg accagtgcac tgagtacatg gccaaggtca ttgagaagct ggtggagcgg 1320
gaggacttcg tggaggcggt gaaggaggag gcagcggctg tggcagcccg gcaggagacg 1380
gactctatcc cgctggtgga cgacatccgc ttccacgtgg ccagcacggt gcagacctac 1440
agcgccatag aggaggcgca gcagcgtctg cgggcactcg aggacctgct cgtgtccatc 1500
ggtctggact gttgagcccc tggctgggca gccccagggg ccaggagctc tcttggagac 1560
aagcatgtgt atgcgtttgt gtgcagctct tcttcctgct ccctgcacat tgagggcttc 1620
atggggggtg cgaggggctc agtggggctt ctcttccctc catgagcctg gagaccccag 1680
gggaggatcc atttgggatg agccccctcc ccccaatgca caagccagcc cccaagaccc 1740
tgggggtgga caccactcag ggaaacctgg ggtgggggtg ggctttggtc ttagcacttt 1800
ccttctccag atccccccta cccaccccag tcccaaatcc agtcctctgg cccttgccta 1860
gccctgaatt gcttctctaa gctggtgttc ccatgcacag ggccattcag gaagggctgg 1920

gggagtgtgt gtggcaataa agcttgaagg caccgtggga gcatgaaaaa aaaaaaaaaa 1980

<210> 84
<211> 1449
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 526155CB1

<400> 84
gcccgccata gagtttagtg gccagagcga ctcttcaggg aggtggcagg aaaggcttgg 60
aacagctgcc ggaggtgacg gagcggcggc cccgcccggt gcgctggagg tcgaagcttc 120
caggtagcgg cccgcagagc ctgacccagg ctctggacat cctgagccca agtcccccac 180
actcagtgca gtgatgagtg cggaagtgaa ggtgacaggg cagaaccagg agcaatttct 240
gctcctagcc aagtcggcca agggggcagc gctggccaca ctcatccatc aggtgctgga 300
ggcccctggt gtctacgtgt ttggagaact gctggacatg cccaatgtta gagagctggc 360
tgagagtgac tttgcctcta ccttccggct gctcacagtg tttgcttatg ggacatacgc 420
tgactactta gctgaagccc ggaatcttcc tccactaaca gaggctcaga agaataagct 480
tcgacacctc tcagttgtca ccctggctgc taaagtaaag tgtatcccat atgcagtgtt 540
gctggaggct cttgccctgc gtaatgtgcg gcagctggaa gaccttgtga ttgaggctgt 600
gtatgctgac gtgcttcgtg gctccctgga ccagcgcaac cagcggctcg aggttgacta 660
cagcatcggg cgggacatcc agcgccagga cctcagtgcc attgcccgaa ccctgcagga 720
atggtgtgtg ggctgtgagg tcgtgctgtc aggcattgag gagcaggtga gccgtgccaa 780
ccaacacaag gagcagcagc tgggcctgaa gcagcagatt gagagtgagg ttgccaacct 840
taaaaaaacc attaaagtta cgacggcagc agcagccgca gccacatctc aggaccctga 900
gcaacacctg actgagctga gggaaccagc tcctggcacc aaccagcgcc agcccagcaa 960
gaaagcctca aagggcaagg ggctccgagg gagcgccaag atttggtcca agtcgaattg 1020
aaagaactgt cgtttcctcc ctggggatgt ggggtcccag ctgcctgcct gcctcttagg 1080
agtcctcaga gagccttctg tgcccctggc cagctgataa tcctaggttc atgacccttc 1140
acctccccta accccaaaca tagatcacac cttctctagg gaggagtcaa atgtaggtca 1200
tgtttttgtt ggtactttct gtttttgtg acttcatgtg ttccattgct ccccgctgcc 1260
atgctctctc ccttgtttcc ttaagagctc agcatctgtc cctgttcatt acatgtcatt 1320
gagtaggtgg gtagccctga tgggggtcgc tctgtctgga gcataaccca caggcgtttt 1380
ttctgccacc ccatccctgc atgcctgatc cccagttcct attaccctac ccctgaccta 1440
ttgaggagc 1449

<210> 85
<211> 1231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 676234CB1

<400> 85
cctcctgctg gacacagaga cacccaccca gcacaccaga cacaccctct gagtcaccta 60
ggccgcctgg ggctgagaag acctaaccga ggggccagat ggcttcgacc ggcttagaac 120
tgctgggcat gaccctggct gtgctgggct ggctggggac cctggtgtcc tgcgccctgc 180
ccctgtggaa ggtgaccgcc ttcatcggca acagcatcgt ggtggcccag gtggtgtggg 240
agggcctgtg gatgtcctgc gtggtgcaga gcacgggcca gatgcagtgc aaggtgtacg 300
actcactgct ggctctgccg caggacctgc aggccgcacg tgccctctgt gtcattgccc 360
tcctgctggc cctgcttggc ctcctggtgg ccatcacagg tgcccagtgt accacgtgtg 420
tggaggacga aggtgccaag gcccgtatcg tgctcaccgc gggggtcatc ctcctcctcg 480
ccggcatcct ggtgctcatc cctgtgtgct ggacggcgca cgccatcatc caggacttct 540
acaacccct ggtggctgag gccctcaagc gggagctggg ggcctccctc tacctgggct 600
gggcggcggc tgcactgctt atgctgggcg gggggctcct ctgctgcacg tgcccccgc 660
cccaggtcga gcggcccgc ggacctcggc tgggctactc catcccctcc cgctcgggtg 720
catctggact ggacaagagg gactacgtgt gaggcggagg tttcccctgg gagcccactg 780
ctcccccactg ccccgccctt tcgaccttgg cctgatgacc agatgccctg ctccatcaca 840
acctccttcc ccaggaaaac ccactttcca aaagcccaag ctacacctgg ctgcagggct 900
gggtcagctg gcctggctga gctcttctca gtggggtccc ctttgatgtt ctccccaag 960
ttgggcagcc tagaggtgtt gggaaccctg gcctgcccc acctccccag taattgtttc 1020
cttccgttgc ccaggacact ggctggcctt ccttctcttc tgagccctcc cctgccccag 1080
gaaccctggc ctcaccaaaa cagcagcagc tcgttggctc caaaaccagg gagcagacca 1140

tgccctccca accctggagt tgtcagggag ggcctgccca tcacctccct ctccccaaca 1200
tccccaccct cgagttggaa ataaagagca t 1231

<210> 86
<211> 858
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 720145CB1

<400> 86
gggccgcggc gaccccgggt ggaggcgacg ttggggttgt ttgggctcct atctttgcct 60
ctggagctca cggtcggcgc ttgtcataac ccagcactag gactgagcgg gcggacgttc 120
gtctttctag aaaggcctag gcgaggccta ggacgagggc ggcggagaac gcggggagaa 180
gcccacccgt gaggagccag ctgccgcgac gcaaatgcag ccctgaggat ttggctactg 240
catataacaa caggggggcaa atcaagtact tcagggttga tttttatgaa gccatggatg 300
actacacatc tgccatagaa gtccaaccca attttgaagt tccatattac aacagagggt 360
tgatactgta taggctggga tattttgatg atgctttgga agatttcaag aaggtcttag 420
acttaaatcc tggatttcaa gatgctactt tgagcttaaa acagactatt ctagacaaag 480
aagaaaaaca aagaagaaat gttgcaaaaa attattgata tttttaactt aatggaagta 540
ttgattcatg atccttacat ctgcatctag ttatcagtaa tttagatatt gagctatttt 600
gatttatatt taagaaatta atacattagc actgaaagtt aaatagtgtg tttaaggtag 660
ttaatttcag gttgaatggg ttttttttaa tgaagtgtaa ataataccaa tgtataagtg 720
tatattatta tattaaatat tatagtaaaa aggaatgtgt ggtattttct tcagcaaaac 780
tatttttgtg attttttat tctcaacttt ttatttaaaa aatgttacat ctgcagaaaa 840
gttgaaagta taataaac 858

<210> 87
<211> 1748
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1001951CB1

<220>
<221> unsure
<222> 65
<223> a, t, c, g, or other

<400> 87
cgcatgcccc gtgcttcgat gtctaagcaa cctgctaact gaggcagcag tggagactgt 60
ggggnggggc aaatgcagct cagagatgag cgttgttgtg gcagccttat ttatccttct 120
gcagttcttt ttccagaaac agcccagtct gctccctgag ggcctttggc tcctcaacaa 180
cctcactgca aacagtccta gtttctgtac ctccttgctc tccctggatc tgattgagcc 240
tctcttacag ctgttgccag tatctaacgt ggtgagcgta atggtgctca cagttctgtg 300
caatgttgca gaaaagggtc ctgcttactg ccagcggctg tggccagggc ccctgcttcc 360
cgccttgctg cacacactag ccttttctga cactgaagta gtaggccaga gtttggaact 420
gctgcaccat caagcatggt gcagcagttc ctgcggcaag cccagcgggg gacagaggaa 480
aaggagagag aggggggctct ggtcagcctt cgtcgaggct tgcagcaccc tgaaacacag 540
caaaccttca tccggagctg tgtctgtata cactgggtaa ccctgatcgt ggagagtgag 600
gctgtgagaa ggcagctcct gccacagggc attgttccag ccttggctgc ctgcatccag 660
tccccccatg tggctgtgct ggaagctctc ggatatgcct tgtcccagct tctacaggct 720
gaggaagctc cagagaagat cattccctcc atcttggcct ccactctccc tcagcacatg 780
ctacaaatgt tgcaacctgg cccaaagctc aaccctgggg tcgctgtgga gtttgcctgg 840
tgccttcatt acatcatctg cagccaggtc agcaatcctc tgctcattgg ccatggggct 900
ctgtctactc tggggttgct gctgttggac ttggctgggg ctgtccagaa aaccgaggat 960
gcaggactgg agctgctggc atgccccgtg cttcgatgtc taagcaacct gctaactgag 1020
gcagcagtgg agactgtggg agggcaaatg cagctcagag atgagcgtgt tgtggcagcc 1080
ttatttatcc ttctgcagtt ctttttccag aaacagccca gtctgctccc tgagggcctt 1140
tggctcctca acaacctcac tgcaaacagt cctagtttct gtacctcctt gctctccctg 1200
gatctgattg agcctctctt acagctgttg ccagtatcta acgtggtgag cgtaatggtg 1260
ctcacagttc tgtgcaatgt tgcagaaaag ggtcctgctt actgccagcg gctgtggcca 1320
gggcccctgc ttcccgcctt gctgcacaca ctagcctttt ctgacactga agtagtaggc 1380

cagagtttgg agctgctgca tctgctgttc ctgtatcagc cagaggctgt tcaggtcttc 1440
ctgcagcagt cagggctgca agcctggaaa aggcatcagg aagaggccca gctccaggat 1500
cgtgtgtatg ctctccagca gacagctctt caagggtgat cttgtttctc aatgtcactc 1560
attcccctct ctcttaacat caagcttgtt ttgtccagta gagcctttgg agatttagga 1620
ccataatgag gtctcatgtt ctctgctccc acacctaagc caagaccttt gggtcccagc 1680
tcctcccccct tccactcagc actatccagg caaggaggac caaaaaggga ctcagtgtgg 1740
tctactta 1748

<210> 88
<211> 4240
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1243349CB1

<400> 88
agctttgcca cagaaagcag tgagcaagaa tgatagctgt ctcttttaaa tgccgttgtc 60
aaattctgag gcgacttact aaagatgaga gtccctacac taaatccgcc agccagacaa 120
agccgcctga tggagcgttg gctgtgagga gacagagcat cccagaggaa ttcaagggct 180
ccacagtcgt cgagctgatg aagaaggaag gcactaccct gggtctgacg gtatcgggag 240
gaattgataa ggatggcaag ccaagagtat ctaatctgcg gcaaggagga attgctgcta 300
gaagtgacca gctggatgtg ggtgactaca tcaaagcagt gaatggaatc aacctggcca 360
aattccgcca tgacgagatc atcagcttgc tgaagaatgt gggagaaaga gtggttcttg 420
aagtagagta cgagcttcca ccggtctctg tgcaaggatc aagtgttatt ttccgaacag 480
tggaggtcac attacataaa gaaggcaata cctttggttt tgtaattcga gggggagcac 540
atgatgatag aaataaatct cgtccagttg tgataacatg tgttcgtcct ggagggcctg 600
ctgacagaga gggcacgatc aaacccggtg acaggttgct cagtgtggat ggaattcggc 660
ttcttggaac tacgcatgct gaagccatga gtattcttaa acaatgtgga caagaagcag 720
cactgctgat agaatatgat gtctcagtaa tggactctgt ggcaacagca tccgggccac 780
tactagtcga agttgccaaa actcctggtg ccagccttgg ggttgcccta actacctcga 840
tgtgctgtaa caaacaagtc attgtcatag acaaaatcaa atctgcaagt attgcagaca 900
gatgtggcgc attgcatgtg ggagatcaca tcctctccat cgatggaacc agcatggagt 960
actgtacact tgcagaagca acccagttcc tggccaacac cactgaccag gtcaagcttg 1020
agatccttcc ccatcatcag acccggctgg ccctaaaggg gcccgaccat gtgaaaattc 1080
agaggagcga caggcaactt acctgggatt cctgggccag caaccacagc agccttcaca 1140
ccaaccatca ctataacacg taccaccctg accattgcag agtaccagcc ctgacattcc 1200
cgaaagcacc tcctccaaac agccctccag ctttggtgtc ttcatccttc tctcctacct 1260
ccatgagtgc atacagcctg agttccctga acatggggac tctacctcga agcctctact 1320
ccaccagccc acgtggaacc atgatgagga ggagactgaa aaagaaagac ttcaaaagct 1380
cattgtcctt agcctccagc acagtaggat tggctgggca ggttgttcac acagaaacca 1440
cagaggttgt tctgacggca gatcctgtca caggatttgg gatccaactg cagggcagtg 1500
tgtttgccac agaaactctc tcttctccac ctctgatttc ctatatcgaa gctgacagcc 1560
cagcagagag atgtggggtg ctacagattg gagacagagt gatggccatc aatggaattc 1620
caacagaaga cagcaccttc gaagaagcca gtcagctcct ccgagactct tcaatcacga 1680
gcaaggtcac actggaaatc gagtttgatg ttgcagagtc tgtcatccca agtagtggaa 1740
catttcatgt aaagctgcct aagaagcaca atgtggaact tggaataacc ataagttcac 1800
catccagtag aaaaccagga gaccccctcg tcatttcaga tatcaagaaa gggagtgtgg 1860
cacacagaac tgggaccctg gaacttgggg ataaattgct cgcaatagat aatatccggc 1920
tggacaactg ttccatggaa gatgcagttc agatcctcca gcaatgtgaa gacctggtga 1980
agctcaaaat ccgcaaagat gaagataatt cagatgagca agaaagttcc ggagcaatta 2040
tttacaccgt ggagcttaaa cgctacgggg ggccccttgg catcacaatt tcaggaactg 2100
aagagccgtt tgatcctata atcatttcaa gcctcactaa aggggggatta gctgaaagaa 2160
ctggcgcaat ccacatagga gaccgaatcc tagccatcaa tagcagcagc ttgaaaggga 2220
agcctctgag tgaagccatc catttgttac agatggcagg agagactgtc accttgaaaa 2280
ttaagaaaca gacagatgcc cagtcagcat cgagccccaa gaagttccct atttctagcc 2340
atttgagtga cctgggggat gtggaggagg actcctcacc agcacagaag ccaggcaagc 2400
tctccgacat gtacccctcc acggtgccca gtgtggacag tgctgtggat tcatgggatg 2460
ggtctgcaat agacaccagc tatggaactg aaggcactag ttttcaggcc tcaggataca 2520
atttcaacac ctatgactgg aggagtccaa aacagagagg cagcttgtcc ccagtcacta 2580
agcctcgaag ccagacttac ccagatgtgg ggctgagtta tgaagactgg gaccggtcca 2640
cagccagtgg ttttgcaggg gctgccgata gtgcagagac agaacaagag gagaacttct 2700
ggtctcaagc gctggaggat ttggaaacct gcggacagtc aggaattctg agagaactgg 2760
aggcaacaat catgtcgggg agcacgatga gtttgaatca tgaggctcca acacctcgca 2820
gtcagctggg gcgacaggcc agcttccagg agcgcagcag ctcgcggccg cactacagcc 2880
aaacaactcg gagcaacacc ctgccttcag atgtgggtag gaagtcagta accctgagaa 2940

aaatgaaaca agaaataaag gagatcatgt ctccaactcc tgtggagctg cacaaggtga 3000
ccttgtacaa ggactctgac atggaggact ttgggttcag tgtagcagat ggcttactgg 3060
agaaaggagt gtatgtcaaa aatattcgcc cagctgggcc aggagatctt ggtggcttaa 3120
agccctatga caggctctta caggtgaatc atgtccgaac cagagacttt gactgctgcc 3180
ttgttgtgcc cctcatagca gaatccggga ataagctgga cctggttatt agtagaaacc 3240
cactggcttc acagaagtct atagaccaac agagtctacc aggagattgg agtgaacaga 3300
acagtgcttt tttccagcag cctagccacg gtggtaattt ggagacacga gaacccacta 3360
atacattata gcaacgcttt ttataaagca ggacaaaaga caatatctac atggtgctaa 3420
aaatccttta agatttctgt gacctttgaa gcacagataa tcaatcaacg tggcattaac 3480
tgcaagcaca ggggtctttt aaatctctct catggctcat gttcacttcc cttttcaagt 3540
tgaagaggtt tctttttttgg tgaccactat ggtatatggt gggcaatgcc ctgccagtcc 3600
caacggtaga gaaaaatagg ccgtcccccca caactctaca attaacatca gaggaaattt 3660
tttacaagtt catcttacta tcacttttta aaaagagaaa catctgtttg aaaatattct 3720
ctgtgatgat ttccttaatt cactttgaaa tcagtttctt actatgaagt cattaatgta 3780
agaacttggc caacaaagct tttcttctca taggctggct ctactagggg aactagtgtt 3840
tggtaaactg ctgggactac cacaatggga ggggtacagg tataaaatta agttatctta 3900
aaatgtttca gcaatgatgc acgtaggaga ccataatagg tggtggtaaa tgttttgccc 3960
acgtatagga atgattttaa ctaagacgta tgctattccc tatgcaacaa attatcaaac 4020
aggatatgtc ttgtgacctg ttttttttt aaggacacat ttttaatagc tgaaaatctc 4080
tgataatgaa ttagagtgtg tagtaaacat gagaattagt tatattatct tatttttaaa 4140
attcaagact aagaacttca gagaatgaag agtctattaa aatgaggttc atcttaatga 4200
taggcaaacc aaactcatac tgcttgacat gttttgaaaa 4240

<210> 89
<211> 2317
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1338201CB1

<400> 89
ataaccctca ctaaagggaa taagcttgcg ccgcagccat gtcccggggg cccgaggagg 60
tgaaccggct cacggagagc acctaccgga atgttatgga acagttcaat cctgggctgc 120
gaaatttaat aaacctgggg aaaaattatg agaaagctgt aaacgctatg atcctggcag 180
gaaaagccta ctacgatgga gtggccaaga tcggtgagat tgccactggg tccccccgtgt 240
caactgaact gggacatgtc ctcatagaga tttcaagtac ccacaagaaa ctcaacgaga 300
gtcttgatga aaattttaaa aaattccaca aagagattat ccatgagctg gagaagaaga 360
tagaacttga cgtgaaatat atgaacgcaa ctctaaaaag ataccaaaca gaacacaaga 420
ataaattaga gtctttggag aaatcccaag ctgagttgaa gaagatcaga aggaaaagcc 480
aaggaagccg aaacgcactc aaatatgaac acaaagaaat tgagtatgtg gagaccgtta 540
cttctcgtca gagtgaaatc cagaaattca ttgcagatgg ttgcaaagag gctctgcttg 600
aagagaagag gcgcttctgc tttctggttg ataagcactg tggctttgca aaccacatac 660
attattatca cttacagtct gcagaactac tgaattccaa gctgcctcgg tggcaggaga 720
cctgtgttga tgccatcaaa gtgccagaga aaatcatgaa tatgatcgaa gaaataaaga 780
ccccagcctc taccccccgtg tctggaactc ctcaggcttc acccatgatc gagagaagca 840
atgtggttag gaaagattac gacaccсttt ctaaatgctc accaaagatg ccccccgctc 900
cttcaggcag agcatatacc agtcccttga tcgatatgtt taataaccca gccacggctg 960
ccccgaattc acaaagggta aataattcaa caggtacttc cgaagatccc agtttacagc 1020
gatcagtttc ggttgcaacg ggactgaaca tgatgaagaa gcagaaagtg aagaccatct 1080
tcccgcacac tgcgggctcc aacaagacct tactcagctt tgcacaggga gatgtcatca 1140
cgctgctcat ccccgaggag aaggatggct ggctctatgg agaacacgac gtgtccaagg 1200
cgaggggttg gttcccgtcg tcgtacacga agttgctgga agaaaatgag acagaagcag 1260
tgaccgtgcc cacgccaagc ccacaccag tgagaagcat cagcaccgtg aacttgtctg 1320
agaatagcag tgttgtcatc cccccacccg actacttgga atgcttatcc atgggggcag 1380
ctgccgacag gagagcagat tcggccagga cgacatccac ctttaaggcc ccagcgtcca 1440
agcccgagac cgcggctcct aacgatgcca acgggactgc aaagccgcct tttctcagcg 1500
gagaaaaccc ctttgccact gtgaaactcc gcccgactgt gacgaatgat cgctcggcac 1560
ccatcattcg atgagaggac agccaaggac tctcccgggc ctctccggtt ctcccttgcg 1620
gaatgatggg cgcatcctgt ctgccacgtg ctgacggtcg ggaagcttca gtggagaggc 1680
ctaactctaa tgtcgcctgc ttaagcaaat catgcttctc tgtttcacgt agttgggttg 1740
acaagtttct gcctttaaga taaatgagta atagtctaat gaccagctca gccatttaaa 1800
atattttctt cctattctgt tcaagaaaca gtaaacttgg tttcaatctt tactgtattt 1860
tttaaatgaa tttttccctt aataacagcc agaataaggg atagtctatg ctttcaggac 1920
tggctttctg cacctgatat gaatgagacc agttttattt tataaagcat gtgctcttaa 1980
tagcattatg tctaaagaag atatcacgta agtttgcatc ttagcatgca aatcataatt 2040

ttaagcaata taaattatga aaatactata taaatgtaat ttaacttaaa atgtttaagt 2100
gtagagcttc cagagatggg ggaaaccccc accctccctc caaccacgcc agagctgtag 2160
gagtgctaag acgctttggc tgcccttatc acagcccacg tagcatcgga ggaacctctc 2220
cgggagctcc tttcctggga gtttggagtt gccacgggtt cggttgaaag ggtcatccgg 2280
gaccagcctg ggggtggagt tcagggcctt ggtttct 2317

<210> 90
<211> 3899
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1405141CB1

<400> 90
cggcacgagc tcgtgccgaa ttcggcacga gaaaccagct gttccgagtc ctcaccagag 60
ccacacagga agaagtcgcg gttaacgtga ctcgggtcat tattcatgtg gttgcccagt 120
gccatgagga aggattggag agccacttga ggtcatatgt taagtacgcg tataaggctg 180
agccatatgt tgcctctgaa tacaagacag tgcatgaaga actgaccaaa tccatgacca 240
cgattctcaa gccttctgcc gatttcctca ccagcaacaa actactgaag tactcatggt 300
ttttctttga tgtactgatc aaatctatgg ctcagcattt gatagagaac tccaaagtta 360
agttgctgcg aaaccagaga tttcctgcat cctatcatca tgcagtggaa accgttgtaa 420
atatgctgat gccacacatc actcagaagt ttcgagataa tccagaggca tctaagaacg 480
cgaatcatag ccttgctgtc ttcatcaaga gatgtttcac cttcatggac aggggctttg 540
tcttcaagca gatcaacaac tacattagct gttttgctcc tggagaccca aagaccctct 600
ttgaatacaa gtttgaattt ctccgtgtag tgtgcaacca tgaacattat attccgttga 660
acttaccaat gccatttgga aaaggcagga ttcaaagata ccaagacctc cagcttgact 720
actcattaac agatgagttc tgcagaaacc acttcttggt gggactgtta ctgagggagg 780
tggggacagc cctccaggag ttccgggagg tccgtctgat cgccatcagt gtgctcaaga 840
acctgctgat aaagcattct tttgatgaca gatatgcttc aaggagccat caggcaagga 900
tagccacccct ctacctgcct ctgtttggtc tgctgattga aaacgtccag cggatcaatg 960
tgagggatgt gtcacccttc cctgtgaacg cgggcatgac tgtgaaggat gaatccctgg 1020
ctctaccagc tgtgaatccg ctggtgacgc cgcagaaggg aagcaccctg gacaacagcc 1080
tgcacaagga cctgctgggc gccatctccg gcattgcttc tccatataca acctcaactc 1140
caaacatcaa cagtgtgaga aatgctgatt cgagaggatc tctcataagc acagattcgg 1200
gtaacagcct tccagaaagg aatagtgaga agagcaattc cctggataag caccaacaaa 1260
gtagcacatt gggaaattcc gtggttcgct gtgataaact tgaccagtct gagattaaga 1320
gcctactgat gtgtttcctc tacatcttaa agagcatgtc tgatgatgct ttgtttacat 1380
attggaacaa ggcttcaaca tctgaactta tggatttttt tacaatatct gaagtctgcc 1440
tgcaccagtt ccagtacatg ggaagcgat acatagccag tgtgagaaag atatcaagtg 1500
tgcttggaat ttctgtagac aatggctatg gccactcgga cgcagatgtt ctgcaccagt 1560
cattacttga agccaacatt gctactgagg tttgcctgac agctctggac acgctttctc 1620
tatttacatt ggcgtttaag aaccagctcc tggccgacca tggacataat cctctcatga 1680
aaaaagtttt tgatgtctac ctgtgttttc ttcaaaaaca tcagtctgaa acggctttaa 1740
aaaatgtctt cactgcctta aggtccttaa tttataagtt tccctcaaca ttctatgaag 1800
ggagagcgga catgtgtgcg gctctgtgtt acgagattct caagtgctgt aactccaagc 1860
tgagctccat caggacggag gcctcccagc tgctctactt cctgatgagg aacaactttg 1920
attacactgg aaagaagtcc tttgtccgga cacatttgca agtcatcata tctgtcagcc 1980
agctgatagc agacgttgtt ggcattgggg gaaccagatt ccagcagtcc ctgtccatca 2040
tcaacaactg tgccaacagt gaccggctta ttaagcacac cagcttctcc tctgatgtga 2100
aggacttaac caaaaggata cgcacggtgc taatggccac cgcccagatg aaggagcatg 2160
agaacgaccc agagatgctg gtggacctcc agtacagcct ggccaaatcc tatgccagca 2220
cgcccgagct caggaagacg tggctcgaca gcatggccag gatccatgtc aaaaatggcg 2280
atctctcaga ggcagcaatg tgctatgtcc acgtaacagc cctagtggca gaatatctca 2340
cacggaaagg cgtgtttaga caaggatgca ccgccttcag ggtcattacc ccaaacatcg 2400
acgaggaggc ctccatgatg gaagacgtgg ggatgcagga tgtccatttc aacgaggatg 2460
tgctgatgga gctccttgag cagtgcgcag atggactctg gaaagccgag cgctacgagc 2520
tcatcgccga catctacaaa cttatcatcc ccatttatga gaagcggagg gatttttgaga 2580
ggctggccca tctgtatgac acgctgcacc gggcctacag caaagtgacc gaggtcatgc 2640
actcgggccg cagcgttctg ggacctact tccgggtagc cttcttcggg cagggattct 2700
ttgaagatga agatggaaag gagtatattt acaaggaacc caaactcaca ccgctgtcgg 2760
aaatttctca gagactcctt aaactgtact cggataaatt tggttctgaa aatgtcaaaa 2820
tgatacagga ttctggcaag gtcaacccta aggatctgga ttctaagtat gcatacatcc 2880
aggtgactca cgtcatcccc ttctttgacg aaaaagagtt gcaagaaagg aaaacagagt 2940
ttgagagatc ccacaacatc cgccgcttca tgtttgagat gccatttacg cagaccggga 3000
agaggcaggg cggggtggaa gagcagtgca aacggcgcac catcctgaca gccatacact 3060

```
gcttccctta tgtgaagaag cgcatccctg tcatgtacca gcaccacact gacctgaacc 3120
ccatcgaggt ggccattgac gagatgagta agaaggtggc ggagctccgg cagctgtgct 3180
cctcggccga ggtggacatg atcaaactgc agctcaaact ccagggcagc gtgagtgttc 3240
aggtcaatgc tggcccacta gcatatgcgc gagctttctt agatgataca aacacaaagc 3300
gatatcctga caataaagtg aagctgctta aggaagtttt caggcaattt gtggaagctt 3360
gcggtcaagc cttagcggta aacgaacgtc tgattaaaga agaccagctc gagtatcagg 3420
aagaaatgaa agccaactac agggaaatgg cgaaggagct ttctgaaatc atgcatgagc 3480
agatctgccc cctggaggac gaagacgagc gtcttaccga attcccttca catcttcaac 3540
gccatcagtg ggactccaac aagcacaatg gttcacggga tgaccagctc gtcttcggtc 3600
gtgtgattac atctcatggc ccgtgtgtgg ggacttgctt tgtcatttgc aaactcagga 3660
tgctttccaa agccaatcac tggggagacc gagcacaggg aggaccaagg ggaaggggag 3720
agaaaggaaa taaagaacaa cgttatttct taacagactt tctataggag ttgtaagaag 3780
gtgcacatat ttttttaaat ctcactggca atattcaaag ttttcattgt gtcttaacaa 3840
aggtgtggta gacactcttg agctggactt agattttatt cttccttgca gagtagtgt 3899
```

<210> 91
<211> 2301
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1686305CB1

<400> 91

```
accgggctcg gcgtgagtcg ctgcggggct gacggggtgg cagtgcggcg ggttacggcc 60
tggtcagacc ataatgactt cagcaaataa agcaatcgaa ttacaactac aagtgaaaca 120
aaatgcagaa gaattacaag actttatgcg ggatttagaa aactgggaaa aagacattaa 180
acaaaaggat atggaactaa gaagacagaa tggtgttcct gaagagaatt tacctcctat 240
tcgaaatggg aattttagga aaaagaagaa aggcaaagct aaagagtctt ccaaaaaaac 300
cagagaggaa aacacaaaaa acaggataaa atcttatgat tatgaggcat gggcaaaact 360
tgatgtggac cgtatccttg atgagcttga caaagacgat agtacccatg agtctctgtc 420
tcaagaatca gagtcggaag aagatgggat tcatgtagat tcacaaaagg ctcttgtttt 480
aaaagaaaag ggcaataaat acttcaaaca aggaaaatat gatgaagcaa ttgactgcta 540
cacaaaaggc atggatgccg atccatataa tcccgtgttg ccaacgaaca gagcgtcagc 600
atattttaga ctgaaaaaat ttgctgttgc tgagtctgat tgtaatttag cagttgcctt 660
gaatagaagt tatacaaagg cttattccag acgaggtgct gctcgatttg ctttgcaaaa 720
attagaagag gccaaaaaag attatgaaag agtattagaa ctagaaccaa ataactttga 780
agcaacaaat gaactcagga aaatcagtca ggctttagca tccaaagaaa actcatatcc 840
aaaggaagct gacatagtga ttaagtcaac agaaggagag cgaaagcaaa ttgaagcaca 900
acagaataag cagcaggcca tttcagagaa agatcggggg aatggatttt tcaaagaggg 960
gaaatatgaa agagcaattg aatgctatac tcgagggata gcagcagatg gtgctaatgc 1020
ccttcttcca gctaacagag ctatggccta tctgaagatt cagaaatatg aagaagctga 1080
aaaagactgc acacaagcca ttttattaga tggctcatat tctaaagctt ttgccagaag 1140
aggaactgca agaacatttt tgggaaagct aaatgaggca aaacaagatt ttgaaactgt 1200
tttacttctg gaacctggaa ataagcaagc agtaactgaa ctctccaaaa ttaaaaagga 1260
attaattgag aaaggacact gggatgatgt ctttcttgat tccacacaaa gacaaaatgt 1320
ggtaaaaccc attgataatc caccgcatcc tggatcaact aaaccactca agaaggttat 1380
tattgaagaa actggtaatt tgatacagac tattgatgtg ccagatagca ctactgctgc 1440
tgctccagag aataatccta ttaatctagc aaatgtaata gcagccacag gcaccacaag 1500
taagaagaat tcaagccaag atgacctttt tcccacaagt gatactccaa gagcaaaagt 1560
attgaaaata gaagaagtca gtgatacttc atccctgcaa cctcaagcca gtttgaagca 1620
ggatgtatgt cagtcttaca gcgagaaaat gcccatagag atagaacaaa aacctgctca 1680
gtttgccaca actgttcttc ctccaattcc tgcaaactcg ttccagctcg aatctgattt 1740
cagacaattg aaaagttctc cagatatgtt gtatcagtat ttaaagcaaa ttgaaccatc 1800
tttgtatcct aagttgtttc agaaaaatct ggatccagat gtattcaacc agatcgttaa 1860
aattctgcat gacttttaca ttgagaaaga aaagccatta ctcatctttg aaatcttaca 1920
aagactttct gaactaaaaa ggtttgatat ggcagtgatg tttatgtcag aaacagagaa 1980
aaagattgca cgtgcattat ttaatcacat agacaagtca ggattgaagg atagttctgt 2040
cgaagaactc aagaaaagat acggtggttg atttccattt ttgctgaaat aattgttttt 2100
gactttcata tgtaaatttt ttctactgaa agtgttttgc tttttaagaa aatgaaatta 2160
tatagcagga aaggactatc tttgaacata agttaattaa ctataaggtg aattgtgatt 2220
taactagtga gaattgtatt caagtgaact ctgtttttct gaaaataaaa atataaacaa 2280
tgagatacaa aaaaaaaaaa a                                           2301
```

<210> 92
<211> 2314

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1688972CB1

<400> 92
ggccctgggg aggccctgag gatagaggag gaggcctggc cgtgcctttc tggagttcac 60
agaccсctaa agacaggagg cagggatggc gtcgtaggcc tggagagcgc ctgggggggcg 120
ggagggcctg aggctgtgct taaggcccag aggtcctgga gcagggcatg tcttgtcctg 180
aggcttaggc agggccccca ggggagttgg ggcaggaggt ggggaggaca ggatcgtgga 240
ggcttcgcat gtcatggtgt ctgttggttt ttaggctttc ttgggaagtc tggtgtttcc 300
tatgtagttt ggggactagg ccacacctct ccctcatgct tgtcaccact gctacacctt 360
gtctgcccac tttctctcct tcccgcccct ggaggtgaga ccctcacctc tctcgtgagc 420
atgctgcctc atgtcttccg tgccgtcgtt ctgtggctgg gtgactgtcc ctcaggccgc 480
caaccctagt tggctctcac agcaacttcc agtgctcctg gacggggact tcctaagcag 540
catgggggcc acacgggccg gccacgcaca ctggcctggg tccgcattcc ctatttcagg 600
ctgttgtctc gggtcagatc tccggaagcg gcaggagcgc tgggtctcct ggtcgtggcc 660
gggcagcgtc tttcttcttc agggtctgga tgccacttcc cgcctcctcc ctgtgcccag 720
tggaaggaga aacatcctag agggtggaga ctgccctgtt tttcttttgc cattgagtcc 780
cagtgtgact caagcacgag gccacctttc tctgccagag gacagaaagc ttcagtgact 840
ggcccagggt cacacagttc tgccgcccag cccagcagcc tgagccagca gtcaggatgg 900
ggtggggaca ctccgtctct cttgagaccc cactcagaga atggggctga cactgtagga 960
gaaatataag caggactcct gctggctgct cccctcctgc atgaggcatg ttagagggac 1020
tgggaccgcc agcccccagc tcccatgccg ggcctcggcc cagtctcagc tcctgtttgg 1080
gaggggggc agagggggcag gggggcagaa gttgagtctg ggagctgggg cagggcacag 1140
gctctagggt ccagctgtcc tcctctcccc tcccctgctg gccactcctg ctgccgtctc 1200
cacagctctg gaatgatgaa gtggacaagg ccgagcagga gctccgcgtg gcccagacag 1260
agtttgaccg gcaagcagaa gtgacccgtc tcttgctgga gggaatcagt agcactcacg 1320
tgaaccacct gcgctgcctc cacgagttcg tcaagtctca gacaacctac tacgcacagt 1380
gctaccgcca catgctggac ttgcagaagc agctgggcag atttcccggc accttcgtgg 1440
gcaccacaga gcccgcctcc ccacccctga gcagcacctc acccaccact gctgcggcca 1500
ctatgcctgt ggtgccctct gtggccagcc tggcccctcc gggggaggcc tcgctctgcc 1560
tggaagaggt ggcccccccct gccagtggga cccgcaaagc tcgggtgctc tatgactacg 1620
aggcagccga cagcagtgag ctggccctgc tggctgatga gctcatcact gtctacagcc 1680
tgcctggcat ggaccctgac tggctcattg gcgagagagg caacaagaag ggcaaggtcc 1740
ctgtcaccta cttggaactg ctcagctagg caggtgcccc catcccccccc gcattctggc 1800
ctaggcagga gaggatgggc gcagcctgcc acttaacttg tttgttggtg acacagttgt 1860
tcagagtggg gagaattcac cccattctgt ccctgcccct agtcacctag ctgtgagggt 1920
gcctgaggct gaatggctcc accccctcccc cagccctgct tctgacctgt ggctctggag 1980
cccctgcccc tgcctgcatc cccgagcacc ccaccctcca ggctccacta aggagggagg 2040
ggctgtctgc agcagctgca ctcagcacct aggccagggt ggggccgccg cagatgggct 2100
caggaagccc caggtgcact cagcgagagc cctgcctttc agttgccaaa agctgcatca 2160
ggggaatgcg gcaaggcaca cagggctctg gcagcccctg gggactgggc gctgcccctg 2220
ggaggggaga gcctggccag ggctggtgtt gggcccggag cagcatcttc cggtgctatc 2280
ctcccctccc acccctcaca gctcaagcca agtc 2314

<210> 93
<211> 1880
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812494CB1

<400> 93
gaagaccctc gggaagcagt cagactgcac tcacctcaga cacactggaa aacagtaaga 60
gaggagagaa agaagcctac tgaggaagaa ataagaaaga tctgcaggga tgaaaaggaa 120
gcgctggggc agaatgagga atctcccaaa cagggtttga ttactatctg gcaactgtca 180
gatctgagtt tctgtcctaa aaatgcactg gcaaattccc tactctcata actgacccat 240
ttccttattt aaacacacac ccacagtgtc tatcttagca attactgaga catgtttagt 300
gacctttcca aatctattat ctcttttgtt atggatgttt ccccagagct aagaaaatct 360
ttgtgtggag cagatgcgtg gatgtgttgc atggatctca catagtttat ctttgattgg 420
atgctgtggt ttgaggctga gggctgcttc catgctttcc tctattgtgc ggcaccctaa 480
ggctactcat acggagagtc cagggacttg ttcttgtctg ggctgagagc ctcatttctg 540

```
catgtgatat aatggactca gctcttgatg tgctactttt atctgggtga caaaataaaa 600
acaatttcat ttcaggcttt cattcttatg catctgcttc taccctctga gtactcccta 660
gatggatttc acatgtcagg tttttcccta ggttcagggt cggagggaga agatggcttt 720
caggtagagt tggagctagt ggagttgact gtggggactc tggatctttg tgagtctgaa 780
gtattgccca agcggaggag gagaaaaagg aataagaagg agaaaagccg agaccaggag 840
gctggggcac atcggactct tctccagcaa actcaagaag aggagccttc cacacagtca 900
tcccaggcag ttgctgcccc cttgggccct ttgctggatg aggccaaagc ccctggtcag 960
ccagagctct ggaatgcact gcttgctgct tgccgagctg gagatgttgg agtgctaaag 1020
ctgcagctag ctcccagccc tgcagaccct agagttctgt ctctgctcag tgccccccttg 1080
ggctccggtg gctttactct cctgcatgca gcagctgcag ctggaagagg ctcagtggtt 1140
cgtctgctgc tggaagcagg tgctgacccc actgtgcagg actctcgggc ccggccacct 1200
tatactgttg cggctgacaa atcaacacgt aatgagttcc gaaggttcat ggagaagaat 1260
ccagatgcct acgattacaa caaggctcag gtgccaggac cattgacacc agaaatggag 1320
gcacggcagg ctacacggaa aagggagcag aaggcagccc ggcggcaacg ggaggaacag 1380
cagcagaggc agcaggagca ggaggagcgt gaaagagaag agcagcggcg atttgccgcc 1440
ctcagtgacc gagagaagag agctctggct gcagagcgcc gactcgctgc ccagttggga 1500
gcccctacct ctccaatccc tgactctgca atcgtcaata ctcgacgctg ctggagttgt 1560
ggggcatccc tccaaggcct gactcccttt cactacctcg acttctcttt ctgctccaca 1620
cgttgcctcc aggatcatcg ccgtcaggca gggaggccct cttcctgatc tcttacagct 1680
ctacctgggg ccaactcagg gacctgagag ggcacattca cagcagccct aggtttttttc 1740
ttccccgtga aaccagagat gatttggaag atgggggtga aggacactcg ggaactaggg 1800
caaagacagg gctagaggta tgtggagctg gtactgtctc tggaatttta atcacaataa 1860
agtttggcaa ggaaaaaaaa                                             1880
```

<210> 94
<211> 879
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2013853CB1

<400> 94

```
ccctctccga ccctttgagc cgtggccgtt gccagatgtc cacaatggga aacgaggcca 60
gttacccggc ggagatgtgc tcccactttg acaatgatga aattaaaagg ctgggcagga 120
ggtttaagaa gttggacttg gacaaatcag ggtctctgag cgtggaggag ttcatgtccc 180
tgccggagct gcgccacaac ccgttggtgc ggcgagtgat cgacgtcttc gacaccgacg 240
gtgatggaga agtggacttc aaggaattca tcctggggac ctcccagttc agcgtcaagg 300
gcgacgagga gcagaagttg aggtttgcgt tcagcattta cgacatggat aaagatggct 360
acatttccaa cggggagctc ttccaggtgc tgaagatgat ggtgggcaac aacctgacgg 420
actggcagct ccagcagctg gtcgacaaaa ccatcatcat cctggacaag gatggcgatg 480
ggaagatatc ctttgaggaa ttcagtgctg tggtcagaga cctggagatc cacaagaagc 540
tggtcctcat cgtatgagcc tttttcttac aagcaccacc caacaacttc tgctttcttc 600
cctatctctt tcaagatttg ctcaagacgt ccaactgtct ctctgactta tctggaagta 660
tttctttttg tgaagccata tgtcctaaca ggagcttcat caccaactca gtgctattaa 720
ttctccttct ctgaatgact cagggtaccc tataggggga agagcaagtc aaatgagcat 780
agtggggaaa gaaaaggaaa tggcttttat aaacatcttt tactttgttt tgattcaaag 840
accaaactag aactttaaaa gttcaaaaat aagaaagta                        879
```

<210> 95
<211> 3162
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2284925CB1

<400> 95

```
tgcgcatcgg gcccatgaga ctgacgcagg accctattca ggttttgctg atctttgcaa 60
aggaagatag tcagagcgat ggcttctggt gggcctgcga cagagctggt tatagatgca 120
atattgctcg gactccagag tcagcccttg aatgctttct tgataagcat catgaaatta 180
ttgtaattga tcatagacaa actcagaact tcgatgcaga agcagtgtgc aggtcgatcc 240
gggccacaaa tccctccgag cacacggtga tcctcgcagt ggtttcgcga gtatcggatg 300
accatgaaga ggcgtcagtc cttcctcttc tccacgcagg tttcaacagg agatttatgg 360
agaatagcag cataattgct tgctataatg aactgattca aatagaacat ggggaagttc 420
```

gctcccagtt caaattacgg gcctgtaatt cagtgtttac agcattagat cactgtcatg 480
aagccataga aataacaagc gatgaccacg tgattcagta tgtcaaccca gccttcgaaa 540
gg~~atg~~atggg ctaccacaaa ggtgagctcc tgggaaaaga actcgctgat ctgcccaaaa 600
gcgataagaa ccgggcagac cttctcgaca ccatcaatac atgcatcaag aagggaaagg 660
agtggcaggg ggtttactat gccagacgga aatccgggga cagcatccaa cagcacgtga 720
agatcacccc agtgattggc caaggaggga aaattaggca ttttgtctcg ctcaagaaac 780
tgtgttgtac cactgacaat aataagcaga ttcacaagat tcatcgtgat tcaggagaca 840
attctcagac agagcctcat tcattcagat ataagaacag gaggaaagag tccattgacg 900
tgaaatcgat atcatctcga ggcagtgatg caccaagcct gcagaatcgt cgctatccgt 960
ccatggcgag gatccactcc atgaccatcg aggctcccat cacaaaggtt ataaatataa 1020
tcaatgcagc ccaagaaaac agcccagtca cagtagcgga agccttggac agagttctag 1080
agattttacg gaccacagaa ctgtactccc ctcagctggg taccaaagat gaagatcccc 1140
acaccagtga tcttgttgga ggcctgatga ctgacggctt gagaagactg tcaggaaacg 1200
agtatgtgtt tactaagaat gtgcaccaga gtcacagtca ccttgcaatg ccaataacca 1260
tcaatgatgt tccccccttgt atctctcaat tacttgataa tgaggagagt tgggacttca 1320
acatctttga attggaagcc attacgcata aaaggccatt ggtttatctg ggcttaaagg 1380
tcttctctcg gtttggagta tgtgagtttt taaactgttc tgaaaccact cttcgggcct 1440
ggttccaagt gatcgaagcc aactaccact cttccaatgc ctaccacaac tccacccatg 1500
ctgccgacgt cctgcacgcc accgctttct ttcttggaaa ggaaagagta aagggaagcc 1560
tcgatcagtt ggatgaggtg gcagccctca ttgctgccac agtccatgac gtggatcacc 1620
cgggaaggac caactctttc ctctgcaatg caggcagtga gcttgctgtg ctctacaatg 1680
acactgctgt tctggagagt caccacaccg ccctggcctt ccagctcacg gtcaaggaca 1740
ccaaatgcaa cattttcaag aatattgaca ggaaccatta tcgaacgctg cgccaggcta 1800
ttattgacat ggttttggca acagagatga caaaacactt tgaacatgtg aataagtttg 1860
tgaacagcat caacaagcca atggcagctg agattgaagg cagcgactgt gaatgcaacc 1920
ctgctgggaa gaacttccct gaaaaccaaa tcctgatcaa acgcatgatg attaagtgtg 1980
ctgacgtggc caacccatgc cgccccttgg acctgtgcat tgaatgggct gggaggatct 2040
ctgaggagta ttttgcacag actgatgaag agaagagaca gggactacct gtggtgatgc 2100
cagtgtttga ccggaatacc tgtagcatcc ccaagtctca gatctctttc attgactact 2160
tcataacaga catgtttgat gcttgggatg cctttgcaca tctgccagcc ctgatgcaac 2220
atttggctga caactacaaa cactggaaga cactagatga cctaaagtgc aaaagtttga 2280
ggcttccatc tgacagctaa agccaagcca cagagggggc ctcttgaccg acaaaggaca 2340
ctgtgaatca cagtagcgta aacaagaggc cttcctttct aatgacaatg acaggtattg 2400
gtgaaggagc taatgtttaa tatttgacct tgaatcattc aagtccccaa atttcattct 2460
tagaaagtta tgttccatga agaaaaatat atgttctttt gaatacttaa tgacagaaca 2520
aatacttggc aaactccttt gctctgctgt catcctgtgt acccttgtca atccatggag 2580
ctggttcact gtaactagca ggccacagga agcaaagcct tggtgcctgt gagctcatct 2640
cccaggatgg tgactaagta gcttagctag tgatcagctc atcctttacc ataaaagtca 2700
tcattgctgt ttagcttgac tgttttcctc aagaacatcg atctgaagga ttcataagga 2760
gcttatctga acagatttat ctaagaaaaa aaaaaaaaga cataaaataa gcgaaacaac 2820
taggaccaaa ttacagataa actagttagc ttcacagcct ctatggctac atggttcttc 2880
tggccgatgg tatgacacct aagttagaac acagccttgg ctggtgggtg ccctctctag 2940
actggtatca gcagcctgtg taacccccttt cctgtaaaag ggttcatct taacaaagtc 3000
atccatgatg agggaaaaag tggcatttca tttttgggga atccatgagc ttcctttatt 3060
tctggctcac agaggcagcc acgaggcact acaccaagta ttatataaaa gccattaaat 3120
ttgaatgccc ttggacaagc ttttcttaaa aaaaaaaaaa aa 3162

<210> 96
<211> 2740
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2376728CB1

<400> 96
cggaccttcg ccttcgctgt cgccgccgcc gccgcccgcg gacgtcgggg ctattagtga 60
aagatggtgg atcgcttggc aaacagtgaa gcaaatacta gacgtataag tatagtggaa 120
aactgttttg gagcagctgg tcaaccttta actatacctg gacgagttct tattggagaa 180
ggagtattga ctaagttgtg caggaaaaag cccaaagcaa ggcagttttt cttgtttaat 240
gatattcttg tatatggcaa tattgtcatc cagaagaaaa aatataacaa acaacatatt 300
attcccctgg aaaatgtcac tattgattcc atcaaagatg agggagactt aaggaatgga 360
tggctaatca agacaccaac taaatctttt gcagtttatg ctgccactgc tacggagaaa 420
tcagaatgga tgaatcatat aaataaatgt gttactgatt tactctccaa aagtgggaag 480
acacccagta atgaacatgc tgctgtctgg gttcctgact ctgaggcaac tgtatgtatg 540
cgttgtcaga aagcaaaatt cacacctgtt aatcgtcgcc accattgccg caaatgtggt 600

tttgttgtct gtgggccctg ctctgaaaag agatttcttc ttcccagcca gtcctctaag 660
cctgtgcgga tttgtgactt ctgctatgac ctgctttctg ctggggacat ggccacatgc 720
cagcctgcta gatcagactc ttacagccag tcattgaagt ctcctttaaa tgatatgtct 780
gatgatgatg acgatgatga tagcagtgac taaggacaca tttgggagta tttaatcagg 840
tgtggctatc tgagaaatca actttggggg aaatgtaaga ttctgagctc tctctctgtt 900
ttgttctagc catgaatttg cctgagaaac ttgtaaccta tgtgcctcaa tatattccat 960
agaaagtagg tccccctgcc ttctcccact cctcacactc ttctacaggg ataggctttt 1020
gcaaatatat cagataaatt ttttgtttct tgtttatttt taggttattt tcttggaagg 1080
ttgggaaaag atgtttgttt taacaggtca tgtactacgt tgttgttttc atttctgtta 1140
taagtaaaac taaaagcaca gaatggtggg aaaggggcta taatgtggtt cattaataat 1200
gttagcagct tttttctaac catcctgtct aatggttaag acaccagtaa caaaaacaca 1260
tgatttggaa atactttggc tttttcatat acctagtggt gccttatcat aatagcactg 1320
ttacatgaaa taagccccta ccttcttact ttctggtttg ttgaaaaaat acactggtgc 1380
tctttgaagt gataaaatga gtgtttatga atgggtgtaa ttaggaaata cttctcatct 1440
gacagctaca aataactaag tttggaggta ttttcactct atatgaataa atattttttcc 1500
ataaaatagt tgtgattata tttttgtttt atataggtcc caaattataa ttgtcaaata 1560
tatattttaa attaataaaa gttgtcattc ttaggaattt ggtttgaaat ttatcagtta 1620
tacagaattg tcatactgca ttagcttcta cctttagtaa gacatatttt ttaggtataa 1680
attcttatgc tttaacatta tttctggatt gaaaatctta taaaaccctt gaaaataaac 1740
agtctctttt ttacaaagcc tgtgttagag cacagattta cctaggcttg aagatttgga 1800
agaaataata tgtaagaatg gcctcaaggc agaccacttt aagtttggct agacttcata 1860
tcgtggaagt attgtctatt tcagtgtgaa actatcttga atttgcaaat atagtgttat 1920
attttataaa gttttgtaaa atcccaaaca atatttctat ttttgtaaaa caattgtatg 1980
tataatctgt atttgaaatc attttgcaat ctatggaaat agagtagcaa ttgctatttc 2040
taaattgtga actttaagtc aatctagatt tattttgaga agtaattgtt cactctttac 2100
ttttgaggca gccattaggt tgaaagtata tatttatcat ataaaacttg atgcgttttg 2160
cactactctt tccatttata tgctgcaaac aactacagtc tttgaaatat ggaaaatcag 2220
cagtctaaag tttgttttaa attctaaatt taaaaaatct tcaaatctga atataccgca 2280
aatgtcatga gaagtttgat tcagtaactt gtgatggagg attctttggt atcttactgt 2340
ttggttaagg cactaatttt acttacctat tagattttga aagtatctga gatatacaaa 2400
tctccctgta ggaaatgtga aagaaaagca caacaaaact agggtttttt gttcatttgc 2460
ttgcttttat gattttttttt ggtttgtttt aatatcaggt ggatttttgt ttctaagcaa 2520
tatatacata aaatcaacca acatatctga aaaggatcat gaaacctgag aaatgctaat 2580
ggagatttgc tggtacatag gaatctagca aattcaggaa ccaaggggaa atgttgtgag 2640
ataacattta cattgtcaac ctttattgac tttgttttta caataaaaaa tattttacaa 2700
cttaaaacgc aaaacgcagt tagaaaagtc tagaaaaaaa 2740

<210> 97
<211> 1079
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2790762CB1

<400> 97
ccggcatgaa gacagactcg cttagtcgcc agtcacttaa gctgagtgca ttgtgatttc 60
caataattga ggcagtggtt ctaaaagctg tctacattaa tgaaaagagc aatgtggcca 120
gcttgactaa gccgccagcg cacagcgcgg caggacgcgc ccgggtctca gcggacttgt 180
gcatgttagc tgtgtagatt tatgtgaggg cttgtaaaac tctggtcttg taaactagtc 240
ttaagcgctt ttaatatgga gacagatgag agcccctctc cgctcccgtg tgggcccgca 300
ggagaagcgg tgatggagag ccgagctcgc cccttccaag cgctgccccg tgagcagtct 360
ccaccacctc ccctgcaaac gtccagtggt gcagaggtaa tggacgttgg ctctggtggt 420
gatggacagt ccgaactccc tgctgaggac cccttcaact tctacggagc ttctcttctc 480
tccaaaggat ccttctctaa gggccgcctc ctcatagacc cgaactgtag tggccacagc 540
ccgcgcaccg cccggcacgc acctgcggtc cggaagttct cccctgacct taagttgctt 600
aaggatgtaa agattagcgt gagctttacc gagagctgca ggagtaagga caggaaggtg 660
ctgtacacag gagcagagcg cgacgtgcgg gcggagtgcg gtctgctcct tagccctgtc 720
agtggggacg tgcatgcttg tccctttggc gggagtgttg gtgacggggt aggcataggg 780
ggtgagagtg ctgataagaa ggatgaggag aatgagctgg atcaggaaaa gagagtggag 840
tatgcagtgc tcgatgagtt agaagatttt actgacaatt tggagctaga tgaagaagga 900
gcaggcgggt tcacggctaa agcaatcgtt cagagagaca gagtggatga agaggccttg 960
aatttcccct acgaggtatg ttggcagccc ctcctctaga gggctcttag caaaacccaa 1020
agagagattt gggaattgca gcatcttttg aaagcaggga aattaaaaaa aaaaaaaaa 1079

<210> 98

<211> 1393
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2869164CB1

<400> 98
ctcggccggc gctgacgcag ccatggcgga ggcggctttg gaagccgtgc ggagcgagtt 60
acgagaattc ccggccgctg caagggagct ctgcgtgcct cttgctgtgc cctacctgga 120
caaacccca actccgctcc acttctaccg ggactgggtc tgccccaaca ggccgtgcat 180
tatccgcaac gctctgcagc actggccggc cctccagaag tggtccctcc cctatttcag 240
agccacagtg ggctccacag aggtgagtgt ggccgtgacc ccagatggtt acgcggatgc 300
cgtgagaggg gatcgcttca tgatgccagc tgagcgccgc ctgcccctga gcttcgtgct 360
ggatgtgctg gagggccggg cccagcaccc tggagtcctc tatgtgcaga agcagtgctc 420
caacctgccc agcgagctgc cccagctgct gcctgatctg gaatcccatg tgccctgggc 480
ctccgaggcc ctgggaaaga tgcccgatgc tgtgaacttc tggctggggg aggcggctgc 540
agtgacttct ttgcacaagg accactatga gaacctctac tgcgtggtct caggagagaa 600
gcatttcctg ttccatccgc ccagcgaccg gcccttcatc ccctatgagc tgtacacgcc 660
ggcaacctac cagctaactg aagagggcac ctttaaggtg gtggatgaag aggccatgga 720
gaaggtgccc tggatcccac tggacccctt ggcgccagac ctagcacggt accctagtta 780
cagtcaggcc caggcccttc gctgcacggt gcgggccggt gagatgctct atctgccggc 840
tctgtggttc caccacgtcc agcagtccca gggctgcatc gcagtgaatt tctggtatga 900
catggaatac gacctcaagt atagttactt ccagctgctc gactccctca ccaaggcttc 960
aggccttgac tgatggagca ctggtgaaca cgaccaagca cgcctcgggg gacggagcca 1020
gcccctccct ggccaggtca attctcgaga gagcctggag tgtgcatgct ggctgctggc 1080
cccgggtcca gcatggcttg agatcagctt tggaggatct tggaatgtgg tcataaggac 1140
tcaaggtgcc aggcaggtct gggtgagggt tctcaggaag ttgccacaca ggtgagcaga 1200
gtggggatca ggtgcagcgg cacctctccc cagcgctgtg atgttgggcg agtcactgcg 1260
tctcgggcat tggtgtcctg tcagtaaaga gataataatg gctgtacctc gcggggctgt 1320
tgtgggcttg gagatgatgt ctatgaggac cagcatggag ctggcacaca ggacatgttg 1380
aataaaaggt agc 1393

<210> 99
<211> 1580
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3317629CB1

<400> 99
ggcctccccg ctttgtgcta cctgcggccc ggctctcctc cgcgccccca ctcgcgctgc 60
gatgacgcgc tcactgttca agggaaactt ttggagtgca gacatcctca gcaccatcgg 120
ctatgacaac attatccaac atctgaacaa tggccgcaag aactgcaaag agtttgaaga 180
ctttctaaaa gaaagggcag caattgaaga gaggtatggc aaagatctgc tcaacctctc 240
taggaagaag ccgtgtggac agtctgaaat caacaccctg aagcgggccc ttgaagtctt 300
caagcagcaa gtagacaatg tggcacaatg tcacattcag cttgcacaga gtttaagaga 360
agaggccagg aagatggaag aattcaggga aaagcaaaaa ctacaacgaa aaaagacaga 420
gctcataatg gatgctatcc ataaacaaaa gagcttacaa ttcaagaaaa ccatggatgc 480
aaagaagaac tatgagcaga aatgccggga caaagatgag gcagaacagg ccgtcagccg 540
gagtgccaac ctggtgaacc cgaagcaaca agaaaagctt tttgtgaaac tggcaacttc 600
aaagaccgca gtagaggact cagacaaagc atacatgctg cacatcggca ccctggataa 660
ggtccgagaa gagtggcaga gtgagcacat caaggcctgc gaggcatttg aggctcaaga 720
atgtgaacga ataaacttct tccggaatgc attgtggtta catgtgaatc agctgtcaca 780
acaatgtgtc accagtgatg aaatgtacga acaagtccga aagagtttag aaatgtgcag 840
cattcagagg gacattgaat actttgtgaa tcaacgcaaa actggacaga ttccaccagc 900
acccatcatg tatgagaatt tctactcctc ccagaagaat gcagtcccag caggaaaggc 960
tacagggcct aacttggcaa ggagaggacc cctcccaatt cctaaaagct caccagatga 1020
tcccaattac tctttggttg atgactacag tttgctctat cagtaaaatc aatgaaacca 1080
gagctttttc cggctagtgc ttctgtgata tggaaagggc acccagagca gcaggaccta 1140
tagccacgtt atgtcagcaa tgaagacttt gaagtgaacc cttgctataa tttttagag 1200
atttaaaatt tatggtagac atttaggaca acataagcaa gtagagttct gcagtttttt 1260
gaagtttaca aattgcccca ttctgaagaa ttattctttc ccagttactc aggttatgaa 1320
tgaattaggt tttcaacatg ggaagcatga aatccacttc tggatttgga gcatccactt 1380

gaggagcaga ggtggcagca gaggattctg agccaccaac tgcagtagtg gctcctttgg 1440
ctttgggcag cctggctgtg gagtttccac ggcgacacac agcctcagtg gtgcaagatt 1500
taaaattacc ttcctttttg gctggaagac ttagaagccg cctgatcata ctttctcatt 1560
ttacagatga ggaaataaac 1580

<210> 100
<211> 840
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3870488CB1

<400> 100
gttaaaaata caaaaatcag ccaggtgtgg tgatggacac ctgtaatccc agctactcgg 60
gagcctgagg taggagaatc gcttgaaccc aggaggcgaa ggttgcagtg agccgagatt 120
gtgcattgca ctccagcctg ggtgacagag caagacttcg tctcaaaaca aaaacacaaa 180
aacaaacaaa aaacagctac cagggtaaag gctgagccgg cgggccctgc cttgactcgg 240
agtcctgact gtccccagat gaagcgccct cagttcagca agccgccagg aggccaccac 300
aagacccagc gctcaggaag gagaaggatc cccagccaca gcagttgcca cccatggacc 360
caaagttgct gaagcagctg aggaaggcag agaaggccga gagggagttc cggaagaagt 420
tcaagtttga aggggagatc gtggttcaca cgaagatgat gatcgacccc aacgctaaga 480
cacgtcgcgg gggtggcaag cacctcggga tccggcgcgg ggagatcctg gaggtgatcg 540
agttcaccag caatgaggag atgctgtgcc gggaccccaa aggcaaatat ggctacgtgc 600
ccagaacagc gctcctgccc ctggagacgg aggtgtacga tgatgtcgac ttctgcgatc 660
ccctggaaaa ccaaccactc cccctgggac ggtaagaccg gtaggcgtgg ggccaggaca 720
gccagccagc ccagcgcccg ctcacccagg agccctggat cccggcgcgg gaaagtcaca 780
gagctgcctg ggcttgtacc tggccacata aagccccagt ttaaagcaaa aaaaaaaaaa 840

<210> 101
<211> 1970
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3886318CB1

<400> 101
gacggaggcc tggcactcgg aagacaacgg atgggagccg tgtgcacgtc gggagctcgg 60
agtgagcgtg agttccgtgc ccaggcccgc gactcggtcc accaggacag cgctccgggt 120
cgacggggtc ctggagccgc gctctgggag ggcgcacgga gggtgaacgg cggcgttagg 180
acccggaggc gcgggcgggg tgggcggcgg ggctaggacc cagcggctcc ggcagagcgg 240
aagcggcggc gggagcttcc gggagggcgg ctcgcaggca ccatgactcc tgtgaggatg 300
cagcactccc tggcaggtca gacctatgcc gtgcccctca tccagccaga cctgcggcga 360
gaggaggccg tccagcagat ggcggatgcc ctgcagtacc tgcagaaggt ctctggagac 420
atcttcagca ggtagagcag agccggagcc aggtgcaggc cattggagag aaggtctcct 480
tggcccaggc caagattgag aagatcaagg gcagcaagaa ggccatcaag gtgttctcca 540
gtgccaagta ccctgctcca gagcgcctgc aggaatatgg ctccatcttc acgggcgccc 600
aggaccctgg cctgcagaga cgcccccgcc acaggatcca gagcaagcac cgcccccctgg 660
acgagcgggc cctgcaggag aagctgaagg actttcctgt gtgcgtgagc accaagccgg 720
agcccgagga cgatgcagaa gagggacttg ggggtcttcc cagcaacatc agctctgtca 780
gctccttgct gctcttcaac accaccgaga acctgtacaa gaagtatgtc ttcctggacc 840
ccctggctgg tgctgtaaca aagacccatg tgatgctggg ggcagagaca gaggagaagc 900
tgtttgatgc cccccttgtcc atcagcaaga gagagcagct ggaacagcag gtcccagaga 960
actacttcta tgtgccagac ctgggccagg tgcctgagat tgatgttcca tcctacctgc 1020
ctgacctgcc cggcattgcc aacgacctca tgtacattgc cgacctgggc ccggcattg 1080
ccccctctgc ccctggcacc attccagaac tgcccacctt ccacactgag gtagccgagc 1140
ctctcaaggc agacctacaa gatggggtac taacaccacc cccaccgccc ccaccaccac 1200
ccccagctcc tgaggtgctg gccagtgcac ccccactccc accctcaacc gcggcccctg 1260
taggccaagg cgccaggcag gacgacagca gcagcagcgc gtctccttca gtccagggag 1320
ctcccaggga agtggttgac ccctccggtg gccgggccac tctgctagag tccatccgcc 1380
aagctggggg catcggcaag gccaagctgc gcagcatgaa ggagcgaaag ctggagaaga 1440
agcagcagaa ggagcaggag caagtgagag ccacgagcca aggtgggcac ttgatgtcgg 1500
atctcttcaa caagctggtc atgaggcgca agggcatctc tgggaaagga cctggggctg 1560
gtgaggggcc cggaggagcc tttgcccgcg tgtcagactc catccctcct ctgccgccac 1620

cgcagcagcc acaggcagag gaggacgagg acgactggga atcctagggg gctccatgac 1680
accttccccc ccagacccag acttgggccg ttgctctgac atggacacag ccaggacaag 1740
ctgctcagac ctacttcctt gggagggggt gacggaacca gcactgtgtg gagaccagct 1800
tcaaggagcg gaaggctggc ttgaggccac acagctgggg cggggacttc tgtctgcctg 1860
tgctccatgg ggggacggct ccacccagcc tgcgccactg tgttcttctc ttaagaggct 1920
tccagagaaa acggcacacc aatcaataaa gaactgagca gaaaaaaaaa 1970

<210> 102
<211> 1258
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4043934CB1

<400> 102
gaatttaata cgatcactat agggaatttg gccctcgagg ccaagaattc ggcacgagga 60
tggggcaatg cttgaggtat cagatgcact gggaggacct ggaagagtac caggccctga 120
ccttcctgac cagaaatgaa attctgtgca tccatgacac cttcctgaag ctctgccctc 180
ctgggaagta ctacaaggag gcaacgctca ccatggacca ggtcagctcc ctgccagctc 240
tgcgggtcaa ccctttcaga gaccgtatct gcagagtgtt ctcccacaaa ggcatgttct 300
cctttgagga tgtgctgggc atggcatctg tgttcagcga gcaggcctgc ccaagcctga 360
agattgagta tgcctttcgc atctatgatt ttaatgagaa tggcttcatt gatgaggagg 420
atctgcagag gatcatcctg cgactgctga acagtgatga catgtctgag gacctcctga 480
tggacctcac gaaccacgtc ctgagtgagt cggatctgga caatgacaac atgctgtcct 540
tctcagagtt tgaacatgca atggccaagt ctccagattt catgtactcc tttcggattc 600
gcttctgggg atgctgatgt agcggcaaat acctgacatg gcagcctcga gggagaccac 660
aggaatcgaa ccccctccag cactggaggg agctggtttg aagtgtgact ttgtactggg 720
cccacactca cctctagaat attgtttatt agataaaaga aaaagctttt ccttagccca 780
tcagatcatc gctttttaaa tgcagggtca tacatggtac tttttattaa gaactgccct 840
ttccagggct tcagtgtgcc agcgatgtca agcaggctgg ggtggcaatc tttctgaggg 900
aatagttcaa atctcaaccc atgtcatagc agggggccaa gccaaatggg atgaaggtcc 960
ctagcaagat acatgtcctt ccctcccttc atcaaaaccc ccgacccccca gcacctcaca 1020
gttcacagct gcacagagac gtgcacatag cagccattcc agccggtgcc cggctcccca 1080
ctccccttct gagggagaag gcatcttggc cctgattgtg ggaccagcaa ttagagttta 1140
tttgctttgc tattcagcca agctctggaa tgaatttggc caccacagac agcctttggg 1200
gtccagctcc tcagggactt tgagacacag aagaagcccc ccattgggtt tttcttat 1258

<210> 103
<211> 685
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4371445CB1

<400> 103
tgagacctgt ctaaatttgg tttgtgcctt aattgagtta gtcattgaac ttccaaaaat 60
gttgctgcct gaagaggaat ttcagtaacg ttcttcatgt ttacaataat tttcccagtt 120
tgtaaaaaca gtctaaatat tttctgaaaa ttgaaaaact tgaaaaacta aaaaattacg 180
ttgctgtaaa ctctgtgggc ggtgtaaaaa gtttagttgt ggaaaaaata catcatttatt 240
ttctttgtct gacatagaag actcagcctg gatttcaaaa ctgaacactg aagtggtttg 300
caacagttac ctggctcttg gcctcaatcc agctggatca aactgagacc tgaaaagaga 360
tattgactcc tgtggcagtg gaaagaccca agaatgccag tcaagaagac agatactgac 420
cgagctttat cattactgga agaatactgc aaaaaattaa ggaaaccaga ggagcagctg 480
ttgaaaaatg cggttaaaaa ggtgatgggt atctttaaga gcagcttatt ccaagctttg 540
ctaggtatgt attatgaaag ttattcatca ttttaatggt cagaattggg cttgatagtt 600
tgaataatgt ggtcaaagaa cctgatgtta accaagtttc aaactgggat aattttcttt 660
aatactggaa gggaagattt gtttt 685

<210> 104
<211> 1886
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5527925CB1

<400> 104
tgctgcgctt ccgcaaagat ggcggcggct gcgggtagct gcgcgcgggt ggcggcctgg 60
ggcggaaaac tgcgacgggg gctcgctgtc agccgacagg ctgtgcggag tcccggcccc 120
ttggcagcgg cagtggccgg cgcggccctg gcaggagcag gagcggcctg gcaccacagc 180
cgcgtcagtg ttgcggcgcg ggatggcagt tttacagtct ccgcacagaa aaatgttgaa 240
catggaataa tatatattgg gaaaccgtct cttcgtaagc agcgcttcat gcagttttct 300
tcactcgaac atgaaggaga atattatatg acaccacgag acttcctctt ctcagtgatg 360
tttgagcaaa tggaacgtaa aacttcagtc aagaagctga caaaaaagga catcgaggat 420
acactgtcag ggatccaaac agctggctgt ggatcaactt ttttcagaga ccttggcgat 480
aaagggctaa tttcatatac cgagtatctt ttcttgctta caatcctcac taaaccccat 540
tctggatttc atgttgcttt taaaatgctg gatacagatg gtaatgagat gattgaaaaa 600
agggaatttt ttaagctgca gaagatcata agtaaacaag atgacttgat gacagtgaaa 660
actaatgaaa ctggatatca ggaagcaata gtgaaagaac ctgaaattaa cacaactctt 720
cagatgcgtt tctttggaaa aagaggacaa agaaaacttc attataaaga atttcgaaga 780
tttatggaaa atttacaaac agagattcaa gaaatggaat tccttcagtt ttctaaaggt 840
ttgagtttca tgagaaaaga agactttgca gagtggctac ttttttttcac taacactgaa 900
aataaagata tttattggaa aaatgtgaga gagaagttgt cagcaggaga gagcattagt 960
ttggatgaat tcaagtcatt ttgccatttt acaacccact tggaagactt tgctattgcc 1020
atgcagatgt tcagtttagc tcatcgtcct gtcagactag cggagtttaa gagagctgtg 1080
aaagtagcaa caggacaaga actctcaaac aatattttgg acactgtctt taagatcttt 1140
gatttggatg gtgatgaatg tcttagtcat gaagagtttc ttggggtgtt aaaaaacaga 1200
atgcatcgag gtttatgggt accacaacat cagagtatac aagaatactg gaagtgtgtg 1260
aagaaagaaa gcattaaagg agtaaaagaa gtctggaaac aagctggaaa aggtcttttt 1320
taataaaaga tataatagta tggcaattat attgttccaa atgtcaaaat ttgtgatttt 1380
ttagaagtac ttgctattta tcttcttaag tcttcattga tattctgtgt gaaataagca 1440
tgtcttgtac ttgctttctg attcataatt ttatgaaaga acttagtaga aagaaaagta 1500
agtataaaaa tagatattgg attctgtcag aaggcctaga tttgaaataa tgttttgtac 1560
ttcggtaaga tggaaaactt agtgattcac tgatttctta gacactctaa tatgatatgc 1620
tttctggaag gataaaacaa atacatatgg gaaaaagtac ttgagaccaa ggccagcatc 1680
aattccagac atcttcatgt tcctaatagg ctaaatgaag ttaaaaactt atttcagatt 1740
tttctcatct gtaccttata tctcataaat ttattgcata ttttatgtca gtagcttagc 1800
tgtttattgt ctttaaaata acatgtaaac ttcaatgttc tatctggaag cagaataaaa 1860
tatttacata gatacaaaaa aaaaaa 1886

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/16636

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 15587 A (DUKE UNIVERSITY) 1 May 1997 (1997-05-01)	1,3, 6-16,18, 19,22, 25-27
A	abstract page 6, line 17 -page 7, line 5 page 16, line 4 -page 32, line 15 page 48 -page 56; examples 5-8 page 66 -page 69; claims 1-25 page 70 -page 72; figures 1A,1B,2 ---	2,4,5,17

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search
12 October 2000

Date of mailing of the international search report
11. 1. 01

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer
Fuchs, U

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL database, Heidelberg, FRG Emest_Hum3 accession number AI083494 18 August 1998 NCI-CGAP: "ox75a07.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1662132 3' similar to TR:Q13802 Q13802 ARGBPIB PROTEIN, mRNA sequence" XP002149761 the whole document ---	11-15
A	WO 97 44347 A (SMITHKLINE BEECHAM CORP., PRESIDENT AND FELLOWS HARVARD COLLEGE) 27 November 1997 (1997-11-27) abstract page 5, line 21 - line 28 page 29, line 21 -page 32, line 28 page 43 -page 46; claims 1-28 ---	1-19,22, 25-27
A	US 5 602 019 A (BEAVO, J.A. ET AL.) 11 February 1997 (1997-02-11) abstract column 7, line 31 - line 67 column 21, line 1 -column 27, line 2 column 127 -column 128; claims 1-7,10-25 -----	1-19,22, 25-27

International application No
PCT/US 00/16636

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Although claim 18 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 20, 21, 23 and 24
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-19, 22, 25-27 partially

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's prote

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20, 21, 23 and 24

Claims 20, 21, 23 and 24 refer to an agonist and an antagonist of a polypeptide of claim 1 without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).
No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the result to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-19, 22, 25-27 partially

Invention 1

An isolated polypetide comprising an amino acid sequence selected from the group consisting of: a) an amino acid sequence having the SEQ ID NO: 1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1, c) a biologically active fragment of SEQ ID NO: 1, d) an immunogenic fragment of SEQ ID NO: 1; an isolated polynucleotide encoding said polypeptide; a recombinant polynucleotide comprising said polynucleotide; a cell transformed with said recombinant polynucleotide; a transgenic organism comprising said recombinant polynucleotide; a method for producing said polypeptide; an isolated antibody which specifically binds to said polypeptide; an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of: a) a polynucleotide sequence having the SEQ ID NO: 53, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to SEQ ID NO: 53, c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); an RNA equivalent of a)-d); a method for detecting a target polynucleotide in a sample having the sequence of said polynucleotide by hybridizing with a probe or by PCR; a pharmaceutical composition comprising an effective amount of said polypeptide; a method for treating a disease or condition associated with decreased expression of functional OXRD, comprising administering to a patient said pharmaceutical composition; a method for screening a compound for effectiveness as an agonist or antagonist of said polypeptide; a method for screening for a compound that specifically binds to said polypeptide or that modulates the activity of said polypeptide; a method for screening a compound for effectiveness in altering expression of a polynucleotide sequence having the SEQ ID NO: 53;

2. Claims: 1-19, 22, 25-27 partially

Invention 2

Idem as subject 1 but limited to SEQ ID NOS: 2 and 54;

3.-50. Claims: 1-19, 22, 25-27 partially

Inventions 3-50

Idem as subject 1 but limited to SEQ ID NOS: 3-31, 34-52 and 55-83, 86-104.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/16636

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9715587 A	01-05-1997	US 5969101 A	19-10-1999
		AU 7520396 A	15-05-1997
WO 9744347 A	27-11-1997	AU 5926296 A	09-12-1997
US 5602019 A	11-02-1997	US 5389527 A	14-02-1995
		US 6015677 A	18-01-2000
		US 5580771 A	03-12-1996
		US 5800987 A	01-09-1998
		US 5789553 A	04-08-1998
		US 5776752 A	07-07-1998
		CA 2085881 A	20-10-1992
		EP 0535216 A	07-04-1993
		WO 9218541 A	29-10-1992